NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Blinatumomab for treating minimal residual B-precursor acute lymphoblastic leukaemia in remission [ID1036]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Blinatumomab for treating minimal residual B-precursor acute lymphoblastic leukaemia in remission [ID1036]

Contents:

Any information supplied to NICE which has been marked as confidential, has been redacted. All personal information has also been redacted.

NICE National Institute for Health and Care Excellence

Blinatumomab for acute lymphoblastic leukaemia for people with minimal residual disease activity in remission [ID1036]

Pre-meeting briefing

This slide set is the pre-meeting briefing for this appraisal. It has been prepared by the technical team with input from the committee lead team and the committee chair. It is sent to the appraisal committee before the committee meeting as part of the committee papers. It summarises:

- the key evidence and views submitted by the company, the consultees and their nominated clinical experts and patient experts and
- the Evidence Review Group (ERG) report

It highlights key issues for discussion at the first appraisal committee meeting and should be read with the full supporting documents for this appraisal

Please note that this document includes information from the ERG before the company has checked the ERG report for factual inaccuracies

The lead team may use, or amend, some of these slides for their presentation at the Committee meeting

Abbreviation	In full
ALL	Acute lymphoblastic leukaemia
ATT	average treatment effect on the treated
CI	Confidence intervals
CHMP	Committee for Medicinal Products for Human Use
CR	Complete remission
DCAS	Direct comparison analysis set
FAS	Full analysis set
HSCT	Haematopoietic stem cell transplantation
ICER	Incremental cost-effectiveness ratio
IPTW	Inverse probability of treatment weighting
LLQ	Lower limit of quantification
MRD	Minimal residual disease
NHS	National Health Service
NMB	Net monetary benefit
OS	Overall survival
PAS	Primary analysis set
PFS	Progression-free survival
PSA	Probabilistic sensitivity analysis
QALY	Quality-adjusted life year
SOC	Standard of care

Key issues: clinical effectiveness

1. Where would blinatumomab (for MRD-pos) fit into NHS practice? Does the modelling reflect this?
2. Is the measurement and definition of ‘MRD’ standardised and available in the NHS? What level is ‘MRD-positive’?
3. Has the prognostic importance of MRD-positivity been clearly established?
4. Is it clear that eliminating MRD is beneficial?
5. Would patients who achieve MRD negativity with blinatumomab always *proceed* to HSCT?
6. What is the most relevant comparator in the marketing authorisation population?
7. Are results of the indirect comparison generalisable for the population in the MA, considering the absence of:
   - (i) patients unable to receive HSCT or tolerate chemotherapy,
   - (ii) patients in second complete remission (CR2)

Key issues: cost effectiveness

1. Are cost-effectiveness results generalisable for the population in the MA, considering the absence of:

   (i) patients unable to receive HSCT or tolerate chemotherapy,

   (ii) patients in second complete remission (CR2)

2. Which parametric curves for OS and RFS are most appropriate for extrapolation?
3. How should cure be modelled? What cure point should be included in the model?

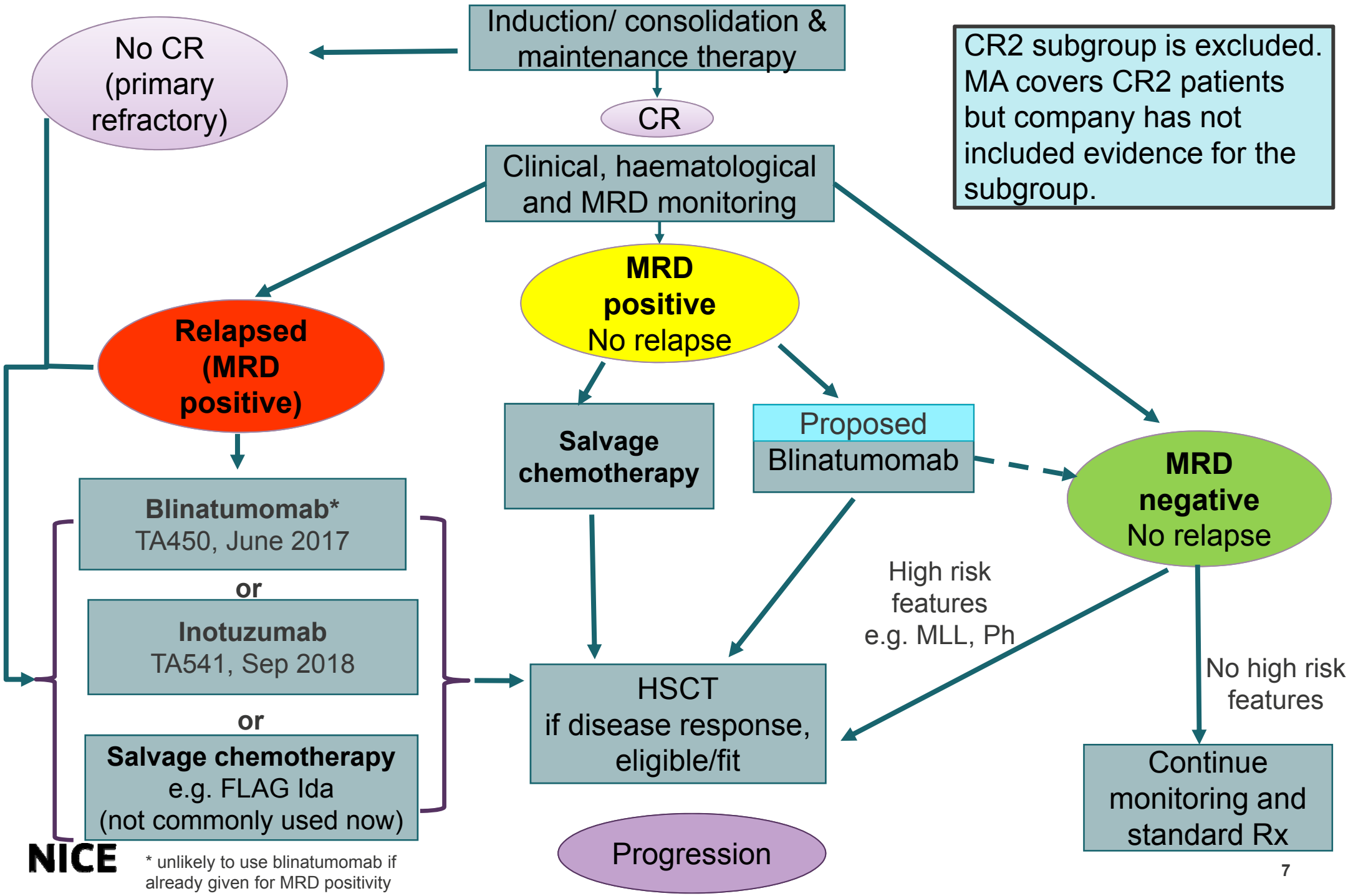
   Company preferred: no fixed cure point; ERG preferred: fixed cure at 5 years in both arms

4. Which post-relapse HRQoL estimate should be used: (i) observed utility of 0.692 among BLAST patients with post-relapse assessment, assumed values of (ii) 0.50 and (iii) 0.25
5. Does the model structure appropriately incorporate HSCT? Should alternative modelling be used or will it have similar uncertainty issues?
6. Which is the most plausible ICER?
7. End of life criteria
8. Equality and innovation
9. Suitable for CDF?

Disease background

- ALL is a rapidly progressing form of cancer of the white blood cells
- Rare in adults - 0.2% of new cancers in UK
- 42% of ALL cases affect adults
- Common in children but children are not covered by the marketing authorisation
- Symptoms include fatigue, breathlessness, infections, bleeding, bruising, fever & sweating. Patients with MRD activity in remission (licensed indication) may not have such extensive symptoms
- 75% of ALL is derived from precursor B-cells (B-cell ALL)
- Most B-cell ALL is Philadelphia chromosome negative (Ph-) (Ph- covered by MA)
- Approximately 44% of adult B-cell ALL patients are expected to relapse and 4% are refractory to available treatments
- MRD: residual ALL present at frequencies below the sensitivity of standard microscopy, but detectable by molecular means in the bone marrow of patients who have met the criteria for haematological complete response.
- No established MRD method for testing, so sensitivity may differ between tests

Treatment pathway for B cell precursor ALL



Blinatumomab (Amgen)

Marketing authorisation	"BLINCYTO is indicated as monotherapy for the treatment of adults with Philadelphia chromosome negative CD19 positive B-precursor ALL in first or second complete remission with minimal residual disease (MRD) greater than or equal to 0.1%." (i.e. $\geq 1 \times 10^{-3}$)
Mechanism of action	Blinatumomab is a T-cell engager targeting CD19 expressed on the surface of cells of B-lineage origin, and the CD3 expressed on the surface of T-cells. It activates endogenous T-cells by connecting CD3 expressed on the T-cell receptor complex with CD19 expressed on benign and malignant B-cells and through this mechanism it harnesses the immune system to kill the cancer cells.
Administration and dosage	It is administered by continuous intravenous infusion using an infusion pump for 28 days, followed by a 14 days treatment free period. Patients may receive 1 cycle of induction treatment followed by 3 additional cycles of consolidation treatment.
List price	The cost of blinatumomab is £2,017 per 38.5 µg vial (list price) The average cost of blinatumomab per cycle at the list price is: £56,476 (28 µg/day for Days 1–28, 28 vials) A simple discount Patient Access Scheme has been approved by NHS England

Decision problem (I)

	NICE scope	Company submission	ERG comments
Population	People with B-cell precursor ALL who have minimal residual disease (MRD) activity while in remission	Adults (≥ 18 years) with Philadelphia chromosome negative and MRD activity B-precursor ALL. Comparative effectiveness and cost-effectiveness evidence is only presented for first complete remission (CR1). The company considers that blinatumomab should be considered in its full marketing authorisation population (including second CR).	2 subgroups were excluded from indirect comparison and economic analysis: (i) patients who are in second haematological remission (CR2) (ii) patients who are unsuitable for HSCT or unable to tolerate chemotherapy.

Decision problem (II)

	NICE final scope	Company submission	ERG comments
Comparator	• Retreatment with combination chemotherapy • Monitor for relapse	• Retreatment with combination chemotherapy Expert opinion suggests that it is highly unlikely that people with MRD activity would only be monitored without any treatment. Therefore monitoring was not considered as a separate comparator, but was incorporated in ongoing chemotherapy regimens.	Blinatumomab may be a treatment option for people who are not eligible for HSCT or cannot tolerate chemotherapy, therefore monitor for relapse should have been included as a comparator for this subgroup
Outcomes	ERG comment: all relevant outcomes included		

Impact on patients – Living with ALL

Submission from Leukaemia CARE

- A rare rapidly progressive disease - most common in a younger population
- Diagnosis with ALL has huge emotional impact, placing a strain on families and friends
- Patients (and their families) experience feelings of:
  - *disbelief, denial, anger, fear, blame, guilt, isolation and depression.*
- Symptoms of active disease include:
  - *fatigue, feeling weak or breathless, sleeping problems, nausea or vomiting, memory loss or loss of concentration, tingling or numbness in extremities, bone or joint pain, bleeding or bruising and infections.*
- Therefore quality of life is affected extensively

Impact on patients – Views on treatments

Submission from Leukaemia CARE

- Patients assessed to be MRD positive following induction treatment, would be considered high-risk, with poor survival (a matter of months)
- There is an urgent need for access to treatments that can induce MRD negativity, prevent relapse and improve survival outcomes.
- Common side effects of blinatumomab:
  - *include fever, headaches, tremors, chills, fatigue, nausea and vomiting.*
- Not unusual for ALL treatment and blinatumomab is generally deemed manageable/tolerable
- In a recent survey, 76% of ALL patients reported that they would be willing to experience additional side-effects for a more effective treatment.
- Potential of outpatient administration is popular with patients
- Use as bridging therapy to stem cell transplant

Professional and clinical expert submissions

Royal College of Pathologists and Sheffield Teaching Hospital NHS Foundation Trust

- Main aim of treatment:
  - Induce remission (clear the majority of the leukaemia)
  - Consolidate remission to reduce relapse (chemotherapy, donor stem cell transplant)
- High unmet need - currently no good treatment of MRD positive patients
- Treatment options are repeating first line chemotherapy (rarely results in long term response) or HSCT, which is often ineffective
- Patients who are MRD positive after chemotherapy have a poor outlook
- Those successfully treated are often young and may go on to live long lives
- Blinatumomab is a safe and effective treatment option, tolerated better than second line chemotherapy
- Clinically meaningful benefits to patients:
  - less patients requiring second line chemotherapy treatment
  - more patients being cured. Increase in length of life more than current care
- Likely to be the only treatment option for people who cannot tolerate chemotherapy
- It can be delivered in outpatient setting

Clinical study evidence: single arm studies

	BLAST (n=116) (Used for economic model)	MT103-202 (n=20)
Design	Phase II, single-arm, open-label, international, multicentre	Phase II, single-arm, open-label, multicentre
Population	• Adult MRD+ BCP-ALL patients in haematological CR after front-line therapy • **Presence of MRD at a level of ≥10^{-3}** • Based in 10 European countries; 7 patients (6.0%) were enrolled in the UK	• Adult MRD+ BCP-ALL patients in haematological CR after front-line therapy • **Presence of MRD at a level of ≥10^{-4}** • 20 patients in Germany received at least one cycle and included in efficacy analysis
Intervention	• Blinatumomab 15 µg/m2/day continuous infusion	• Blinatumomab 15 µg/m^2/day continuous infusion
Primary outcome	• Proportion of patients with complete MRD response	• MRD response rate within 4 treatment cycles
Key secondary outcomes	• RFS at 18 months post initiation • OS; HRQoL	• MRD response after any cycle • MRD progression

ERG comments: Single-arm studies lead to performance bias, detection bias and selection bias.

Patient characteristics BLAST and MT103-202

Baseline characteristic	BLAST (n=116)	MT103-202 (n=20)
Male sex, n (%)	[redacted]	[redacted]
Median age (range), years	[redacted]	Mean age: [redacted]
Relapse history, n (%)		
First CR	[redacted]	NR
Second CR	[redacted]	NR
Third CR	[redacted]	NR
Baseline MRD levels, n (%)	[redacted]	NR
≥10^{-1} <1		
≥10^{-2} <10^{-1}	[redacted]	NR
≥10^{-3} <10^{-2}	[redacted]	NR
<10^{-3}	[redacted]	NR
Below LLQ or Unknown	[redacted]	NR
Philadelphia chromosome disease status, n (%)	Positive [redacted]	Positive [redacted]
	Negative [redacted]	Negative [redacted]

ERG comments: Majority of BLAST study patients (84%) had a baseline MRD level between 10^{-3} and 10^{-1}, where patients are classed as MRD+ when measurable to 10^{-4}. BLAST MRD levels may not necessarily reflect those of the UK population, but reflect the eligibility criteria for the blinatumomab studies.
Note: Red boxes indicate focus of model

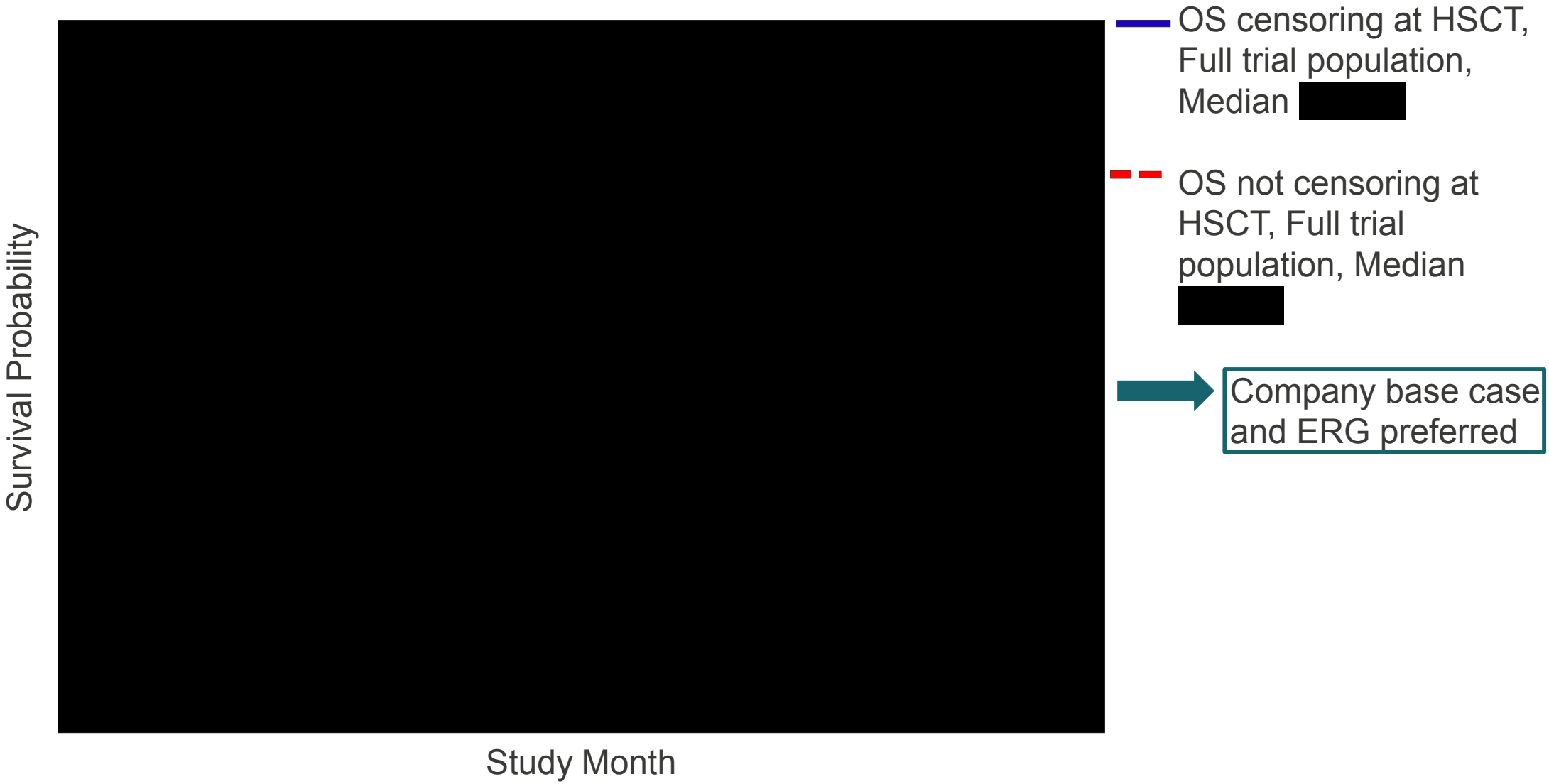
OS and RFS outcomes in BLAST and MT103-202

Study	BLAST (n=116)		MT103-202 (n=20)
Outcome	OS/RFS not censored at HCST (CS primary analysis)	OS/RFS censored at HCST	NR
OS outcomes			
Events, n (%)	[redacted]	[redacted]	NR
Censors, n (%)	[redacted]	[redacted]	NR
OS % at 18 months, (95% CI)	[redacted]	[redacted]	NR
Median (months)	[redacted]	NR	NR
RFS outcomes			
Events, n (%)	[redacted]	[redacted]	NR
Censors, n (%)	[redacted]	[redacted]	NR
RFS % , (95% CI)	[redacted]	[redacted]	[redacted]
95% CI	[redacted]	[redacted]	
Median RFS(months)	[redacted]	[redacted]	[redacted]

NICE

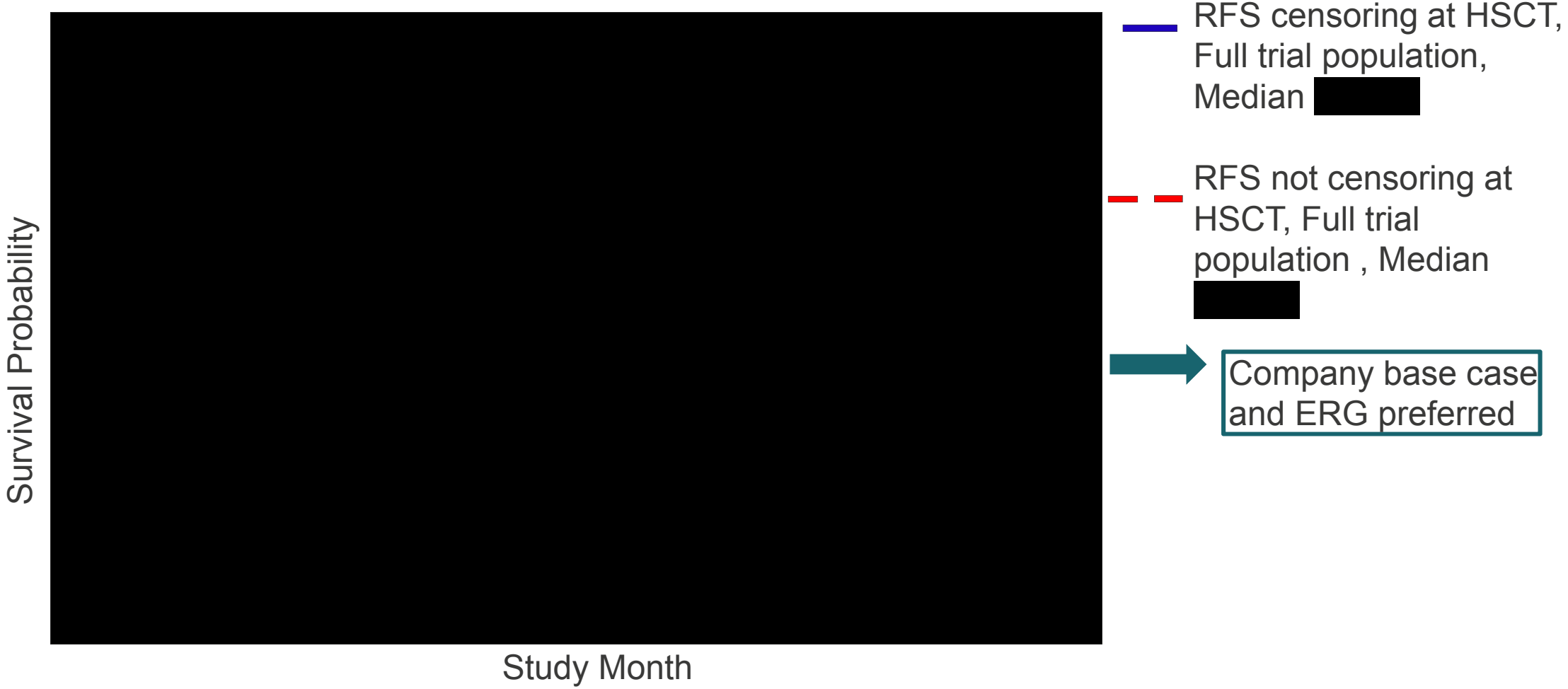
ERG considers OS/RFS **not** censored at HSCT most appropriate (company's base case).

Overall survival results BLAST, Full trial population



- Study MT103-202 did not include OS as an outcome measure

Relapse free survival results BLAST, Full trial population



[redacted]

Results: MRD response and QoL in BLAST and MT103-202

	BLAST ([redacted])	MT103-202 (n=20)
Patients with complete MRD response after 1 cycle, n (%, 95% CI)	[redacted]	[redacted]
Patients with complete MRD response after ≥1 cycle, n (%, 95% CI)	[redacted]	[redacted]
Duration of median MRD response, months	without censoring [redacted] with censoring [redacted]	[redacted]

ERG comments: There was a higher rate of response for patients in CR1 82% (95% CI 72% to 90%), than in CR2 71% (95% CI 54% to 85%) or CR3 50% (95% CI 1% to 99%); but, only 2 patients in CR3. Hence results on subgroup should be treated with caution

- No significant difference for other subgroup analyses

- **EORTC QLQ-30:** Outcomes indicated some [redacted] in HRQoL, [redacted]. By the end of the BLAST study, [redacted]
- **EQ-5D:** Results did not change significantly by the end of the BLAST study

Comparative effectiveness vs chemotherapy

Comparator

- Data on the effectiveness of chemotherapy came from a historical control Study 20120148
- Covers blinatumomab MA population
- Exclusion criteria: use of blinatumomab within 18 months of MRD detection
- Primary endpoint: haematological RFS; secondary endpoints: OS, mortality rate
- Historical study subgroup of the population used in propensity score model to adjust for differences with BLAST population

BLAST subgroup and historical study subgroup are trimmed to match each other according to the following criteria:

- Ph- BCP- ALL;
- First complete haematological remission (CR1);
- MRD+ at a level of ≥1 x 10^{-3};
- ≥18 years old at MRD positivity (historical comparator) or first blinatumomab treatment (BLAST);
- Complete baseline covariate set;
- Time to relapse greater than 14 days from MRD detection (applied to historical study);
- Excludes patients in CR2 and CR3 because comparator doesn't cover them
- Trimming resulted in BLAST subgroup of [redacted] patients and historical study subgroup of [redacted] patients

Comparative effectiveness vs chemotherapy: Inverse probability of treatment weighting

- Due to differences between the populations of BLAST and the historical control study, comparative analyses were undertaken using subsets of the original study populations which were restricted to patients with Ph- disease in CR1 only: BLAST subgroup [n=[redacted]] and historical control [n=[redacted]]
- A propensity score model was constructed and used to generate weights which were applied to the historical control, with the aim of approximating the response to standard care chemotherapy that would be expected in a population with the same characteristics as the BLAST subgroup
- The resulting average treatment effect on the treated (ATT) estimates are applicable to Ph- and CR1 individuals only. This analysis suggested a hazard ratio (HR) [redacted]

ERG comments

- Method used by company is appropriate given limited data set
- Results are representative only of the CR1 population (narrower than MA)
- HSCT unobserved confounders: HSCT rate in BLAST (76%) is higher than the historical control study (37%)
- Limitations to non-randomised data: not possible to account for unobserved confounders and not clear if uncertainty surrounding the method use was accounted for
- Reported treatment effects likely to underestimate associated uncertainty – to be interpreted with caution
- Lack of clarity: stabilised weights presented in clinical effectiveness section, while standard (non-stabilised) weights used in economic model but clarified with company that there is no impact

OS results from propensity score method: BLAST subgroup [redacted] and historical study subgroup [redacted]

Blinatumomab [redacted]

Historical study subgroup [redacted]

Outcome	Median (months)		HR (95% CI)
	Standard care	Blinatumomab	Primary analysis
OS	[redacted]	[redacted]	[redacted]

RFS results from propensity score method:

BLAST subgroup [redacted] and historical study subgroup [redacted]

Outcome	Median (months)		HR (95% CI)
	Standard care	Blinatumomab	Primary analysis
RFS	[redacted]	[redacted]	[redacted]

Adverse events: Safety analysis

Pooled data from BLAST (n=116) and MT103-202 (n=20)

Event	Treatment-emergent AEs [redacted]	Treatment-related AEs [redacted]
All AEs, n (%)	[redacted]	[redacted]
Serious	[redacted]	[redacted]
Grade ≥3	[redacted]	[redacted]
Grade ≥4	[redacted]	[redacted]
Fatal (occur within 30 days of blinatumomab treatment)	[redacted]	[redacted]
Leading to permanent discontinuation of blinatumomab	[redacted]	[redacted]
Serious	[redacted]	[redacted]
Grade ≥3	[redacted]	[redacted]
Grade ≥4	[redacted]	[redacted]
Fatal	[redacted]	[redacted]

- Events occurred in more than 20% of patients: [redacted] The most common treatment emergent AEs of blinatumomab were: [redacted]
  All patients experienced at least one treatment-emergent AE.
- Data included in economic model

Summary of ERG's comments on clinical evidence

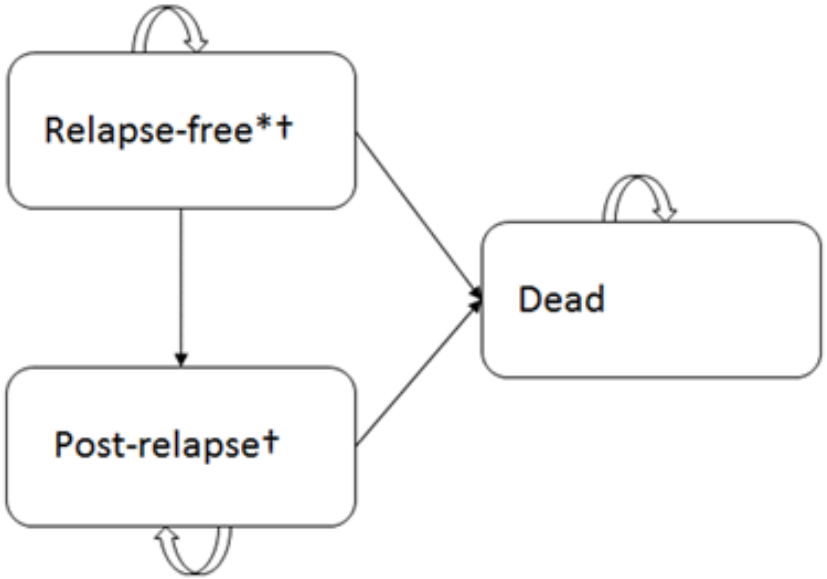
Key areas of uncertainty:

- Only single-arm studies – these were well conducted but subject to inherent bias
- Absence of clinical evidence subgroups excluded from the comparative analysis (patients with CR2+)
- Generalisability to the full population in NICE scope and MA: the treatment effect estimates reflect a narrower population than NICE scope
- Excluded comparator: monitoring for relapse for a subgroup of patients unable to undergo HSCT or tolerate chemotherapy: unclear whether any relevant comparator data exist
- Treatment effects (HR) ignore uncertainty around estimated propensity score weights, and therefore it is likely that estimates underestimate the total uncertainty of the reported HR, resulting in erroneously narrow confidence intervals. HR results should be interpreted with caution.

Cost effectiveness

Company's economic model: structure

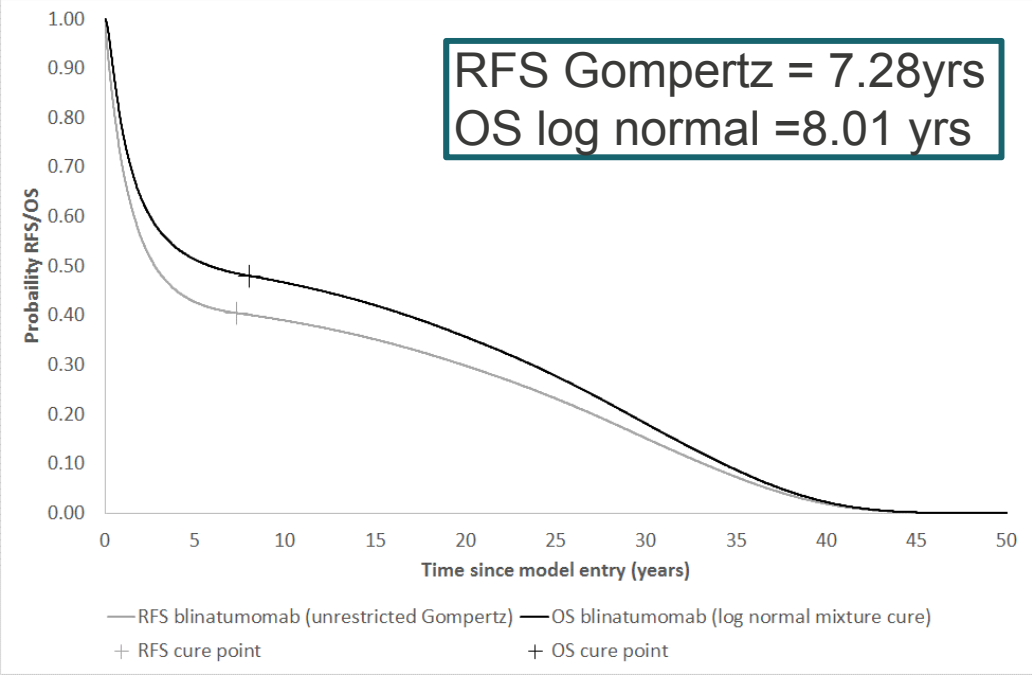
- Partitioned survival model based on RFS and OS. This structure does not allow for tracking of HSCT either before or after relapse.
- The principal benefits of HSCT in avoiding/delaying relapse are implicitly accounted for in the RFS and OS outcomes.
- The QALY losses and costs associated with the HSCT procedure and post-HSCT survival are reflected within two HSCT sub-models applied to the main partition survival structure. The pre-relapse HSCT sub-model is not causally related to RFS or OS, whilst the post-relapse HSCT sub-model is partially related to RFS.



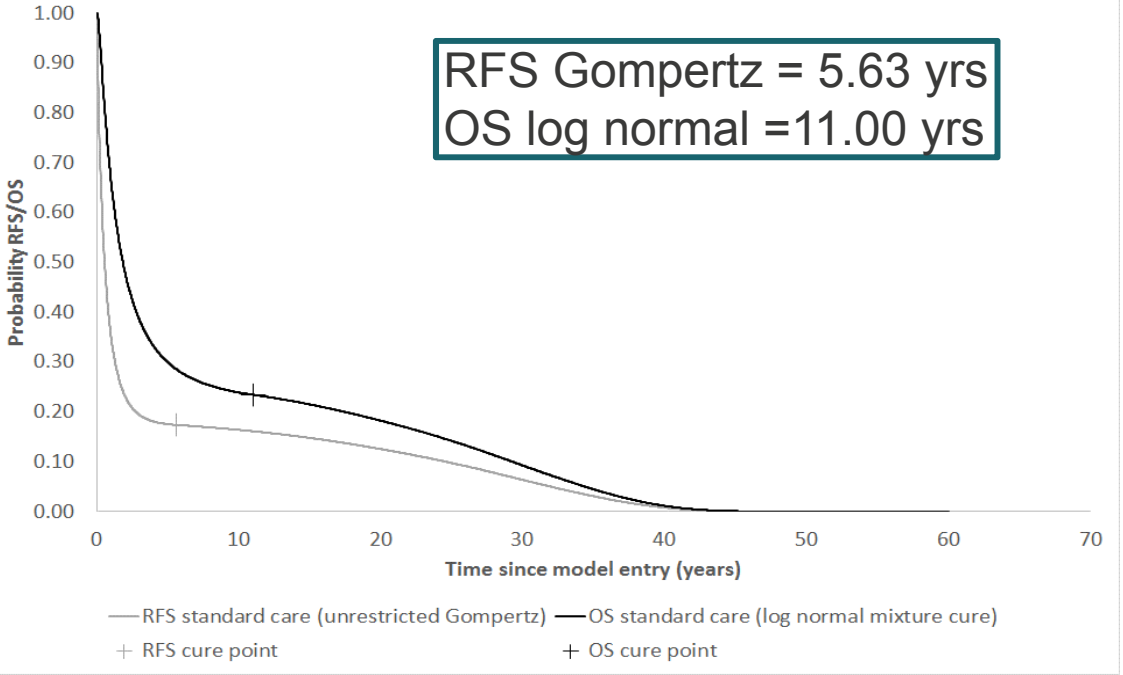
** RFS time divided into time on treatment and post-discontinuation*
† Patients may enter state-specific HSCT sub-model

Company's economic model: RFS/OS and cure point

Company base case: RFS and OS cure points blinatumomab arm



Company base case: RFS and OS cure points SoC arm



- RFS is based on a parametric (Gompertz) model fitted to the treatment-specific RFS time-to-event data
- OS is modelled using a parametric (log normal) mixture cure model fitted to the OS time-to-event data
- Distributions in company's model (RFS and OS) chosen based on a subset of models with best fit and good BIC
- Cure fraction is predicted by model and not fixed in time. Leads to different time points for cure as graphs show

NICE

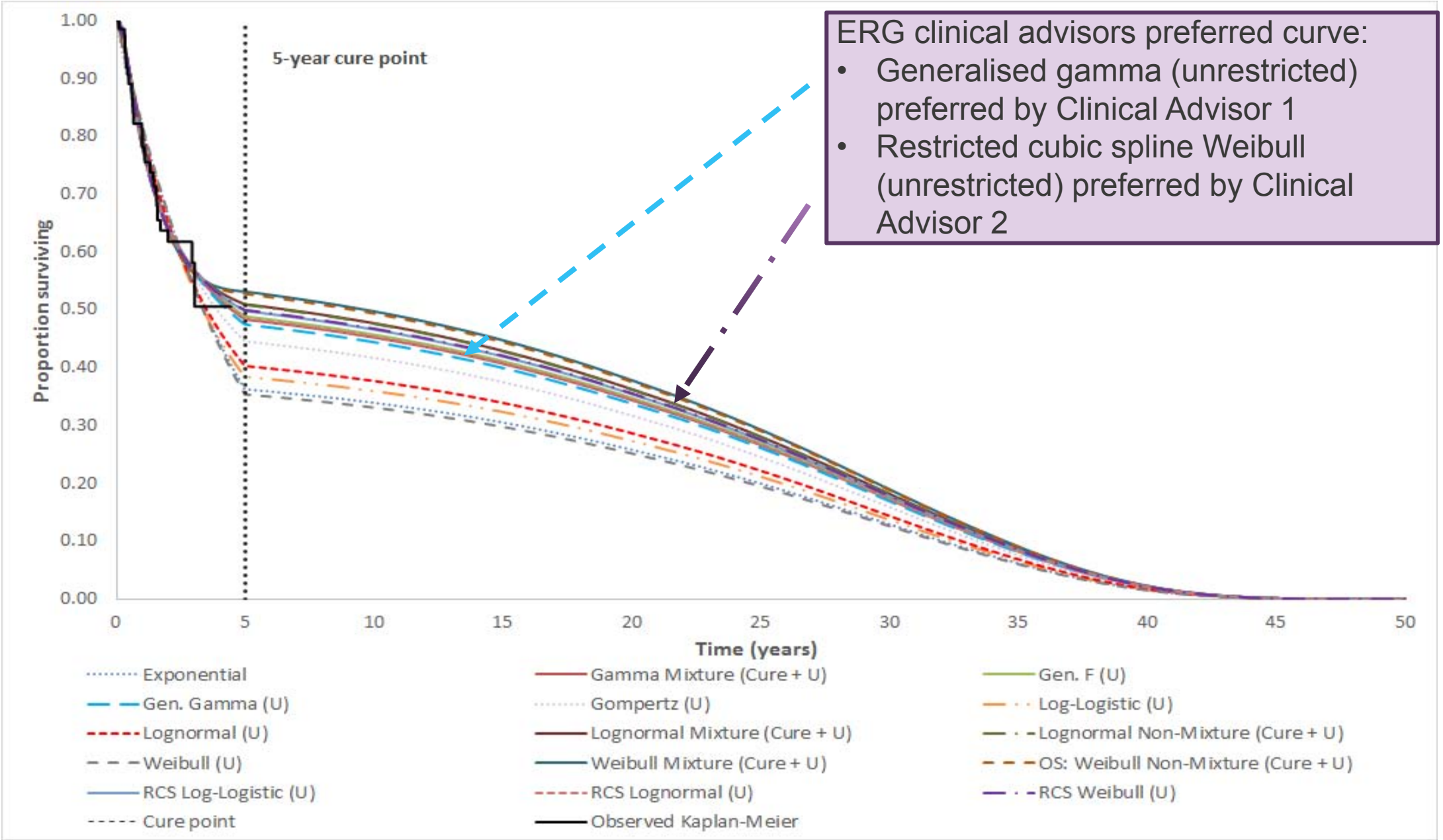
ERG comment: it is not clinically plausible to apply models which feature such a large gap between those achieving cure pre- and post-relapse

ERG critique of company model structure

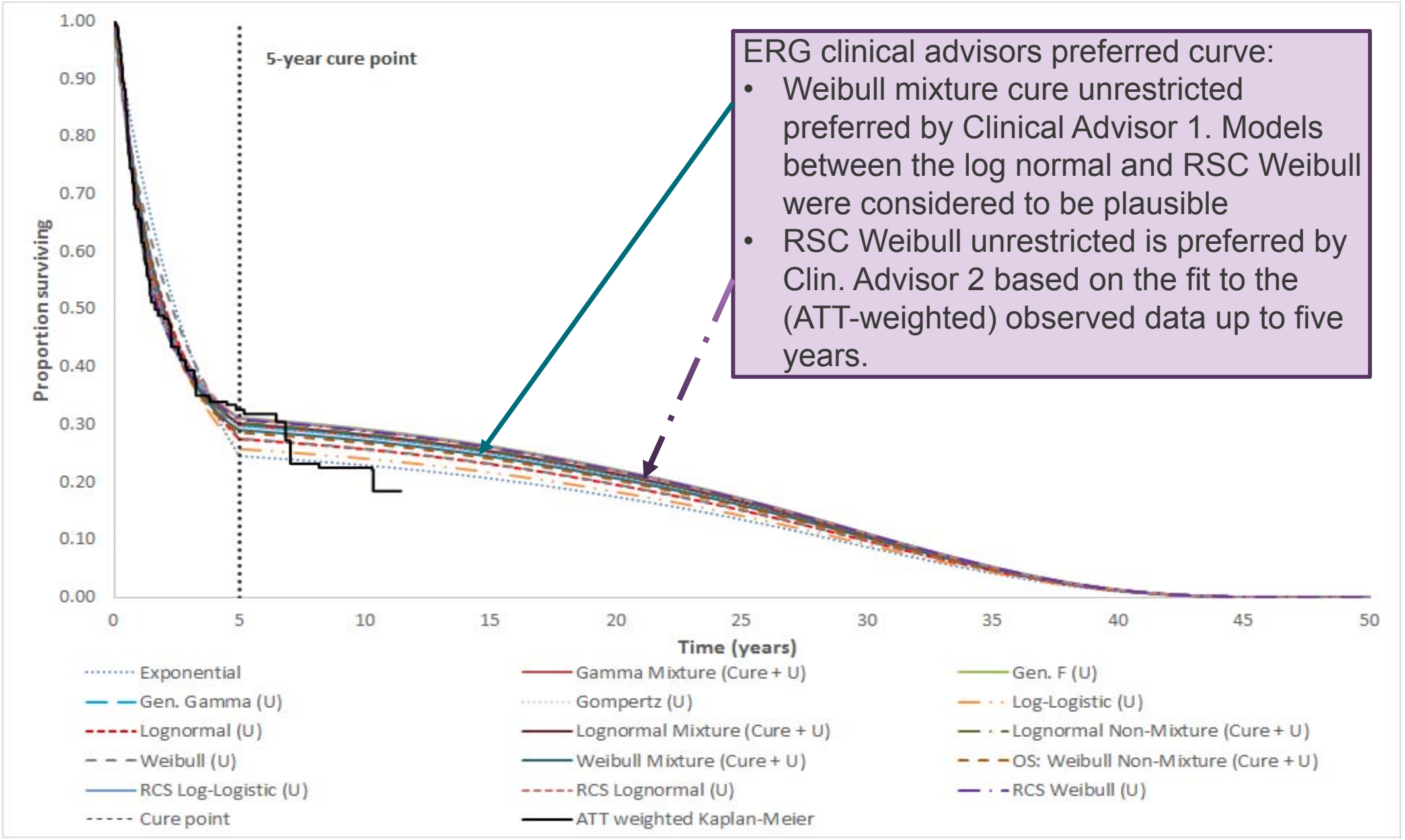
- Cure model appropriate: patient is considered cured if no relapse within 5-years
- Not clear if selected subset of RFS/OS models are clinically plausible (some do not predict a cure fraction and are inappropriate)
- Cure points with different time points (as company base-case) result in large gaps between cure pre- and post-relapse which is not clinically plausible
- Uncertainty regarding proportion of RFS deaths - decreasing them in the blinatumomab group leads to a less favourable ICER
- ERG think more appropriate to apply fixed cure at 5 years and prefers cure unrestricted model
- Model structure not appropriate for tracking HSCT due to:
    a. absence of causal link between HSCT uptake and its impact on RFS and OS outcomes;
    b. model does not estimate probability of receiving HSCT (cannot track patients who undergo HSCT post-relapse);
    c. adoption of questionable assumptions regarding HSCT receipt : BUT no substantial impact on ICER
    d. likely underestimation of post-HSCT costs – relied on survival data only for transplanted cohort; ERG testing shows that increasing post-HSCT costs and HRQoL decrements leads to increased ICER for blinatumomab vs SoC . Still not a big impact on ICER
    e. ERG suggests alternative model (eg. semi-Markov) to fully capture HSCT use

Although, ERG has explored alternative assumptions and models, it notes that data to populate transitions for other models may be limited and may be subject to selection bias and uncertainty.

ERG exploratory and preferred OS for blinatumomab



ERG exploratory and preferred OS for SoC



ERG clinical advisors preferred curve:

- Weibull mixture cure unrestricted preferred by Clinical Advisor 1. Models between the log normal and RSC Weibull were considered to be plausible
- RSC Weibull unrestricted is preferred by Clin. Advisor 2 based on the fit to the (ATT-weighted) observed data up to five years.

Company's model inputs and ERG comments

	Company	ERG comments
Population	Ph- MRD+ BCP-ALL in CR1 (MRD+ ≥1 x 10-3) (BLAST subgroup & historical comparator subgroup with ATT weights)	• Reflects patients likely to tolerate chemotherapy • Narrower than MA as it excludes CR2 patients (due to lack of data) • Cannot assess the cost-effectiveness of blinatumomab in these excluded population groups.
Comparator	SoC - chemotherapy regimen	• SoC chemotherapy regimen comprised of vincristine, prednisolone, mercaptopurine, methotrexate and prophylaxis against CNS relapse using intrathecal methotrexate (treatment up to 2 years) • Excludes "monitor for relapse" (may be relevant to patients unable to undergo HSCT or tolerate chemotherapy)
Costs	Active treatment costs (inpatient and out-patient setting, blinatumomab, HSCT, salvage chemotherapy)	No major issues with cost inputs
Dataset	BLAST subgroup and historical study subgroup with ATT weights	• Mean number of blinatumomab cycles in BLAST= 1.86 • IPTW propensity score methods appropriate given the absence of RCT evidence but introduce uncertainty

Company's model inputs: Utility values

Health state	Utility
Relapse-free utility [Blinatumomab, on-treatment, >6 months prior to death, cycle 1†; cycle 2+†]	0.792; 0.832
Relapse-free utility [Blinatumomab, off-treatment, >6 months prior to death, cycle 1†; cycle 2+†]	0.802; 0.842
SoC, relapse-free, >6 months prior to death	0.806
Post-relapse utility [Blinatumomab and SoC, >6 mos prior to death	0.692
General population utility decrement*	-0.02
HSCT utility decrement [1-12; 13-24; 25-60; 61+ months]	-0.170; -0.010; -0.020; 0.000

ERG concerns regarding plausibility of HRQoL estimates

- unrealistically high post-relapse utility estimate of 0.692 (per ERG clinical expert opinion)
- ERG ran exploratory analysis 7 and applied alternative post-relapse utility estimates (observed utility of BLAST patients with post-relapse assessment 0.819, assumed values of 0.50 and 0.25)
- Results show only a minor impact on ICER

Cost effectiveness results: company's base case (post-clarification submission, PAS included)

Option	QALYs	Costs	Inc. QALYs	Inc. Costs	ICER
Probabilistic results (company's base case post clarification: unrestricted Gompertz function for RFS, log normal mixture cure model for OS, not-fixed cure point predicted by model)					
Blinatumomab	7.11	[redacted]	2.92	£83,634	£28,655
Standard care	4.19	[redacted]	-	-	-
Deterministic results (company's base case post clarification)					
Blinatumomab	7.23	[redacted]	3.02	£83,800	£27,779
Standard care	4.21	[redacted]	-	-	-

Company's updated model submitted post-clarification with the following amendments: (i) maximum annual mortality risk capped at 100%; (ii) pump costs included for all days after the first inpatient stay; (iii) general population utilities based on Ara and Brazier (2010), and (iv) post-relapse allogeneic HSCT not initiated after 5 years

Cost breakdown: company's base case

Cost Category	Blinatumomab (£)	SOC (£)	Incremental (£)
Pre-Relapse			
Blinatumomab and SOC maintenance treatment			
Medication	[redacted]	[redacted]	[redacted]
Administration			
Hospitalisation	[redacted]	N/A	[redacted]
Outpatient visits	[redacted]	[redacted]	[redacted]
Infusion pump	[redacted]	N/A	[redacted]
Total medication and admin.	[redacted]	[redacted]	[redacted]
Allo-SCT	[redacted]	[redacted]	[redacted]
Other inpatient	[redacted]	[redacted]	[redacted]
Other outpatient	[redacted]	[redacted]	[redacted]
Total pre-relapse	[redacted]	[redacted]	[redacted]
Post-relapse			
Salvage therapy	[redacted]	[redacted]	[redacted]
Allo-SCT	[redacted]	[redacted]	[redacted]
Other inpatient	[redacted]	[redacted]	[redacted]
Other outpatient	[redacted]	[redacted]	[redacted]
Total post-relapse	[redacted]	[redacted]	[redacted]
Terminal care	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]

Note: Mean number of blinatumomab cycles in BLAST= 1.86

Cost effectiveness results: ERG's corrected version of company's base case (PAS incl.)

Option	QALYs	Costs	Inc. QALYs	Inc. Costs	ICER
Deterministic results (company's base case post clarification, used by ERG for expl. analyses)					
Blinatumomab	7.23	[redacted]	3.02	£83,800	£27,779
Standard care	4.21	[redacted]	-	-	-
ERG's rebuilt deterministic model (exploratory analysis 1: minor errors corrected)					
Blinatumomab	7.21	[redacted]	3.00	£83,264	£27,717
Standard care	4.21	[redacted]	-	-	-

ERG comment: PSA cost-effectiveness based on company's probabilistic model:

- Approx. 80% of ICER estimates lie below the £50,000/QALY threshold and 50% below the £30,000/QALY threshold.

ERG exploratory analyses results (I)
(deterministic results, PAS included)

Option	QALYs	Costs	Inc. QALYs	Inc. Costs	ICER
Company's base case deterministic version: RFS Gompertz (U) & OS lognormal mix cure					
Blinatumomab	7.23	[redacted]	3.02	£83,800	£27,779
Standard care	4.21	[redacted]	-	-	-
ERG exploratory analysis 1 – Correction of errors identified during model verification					
Blinatumomab	7.21	[redacted]	3.00	£83,264	£27,717
Standard care	4.21	[redacted]	-	-	-
ERG exploratory analysis 2 – Fixed cure point applied to all surviving patients at 5 years					
Blinatumomab	7.37	[redacted]	2.77	£83,803	£30,304
Standard care	4.61	[redacted]	-	-	-
ERG exploratory analysis 3 –Analyses 1 and 2 combined (ERG-preferred model)					
Blinatumomab	7.35	[redacted]	2.75	£83,268	£30,227
Standard care	4.59	[redacted]	-	-	-

- **ERG comment analysis 2**: 5-year cure point is applied to original model, hazard of death is switched to the general population at year 5 and beyond.
- **ERG comment analysis 3**: ERG's preferred model is company's updated model with corrected errors and added 5-year fixed cure point. The uncertainty based on original parametric RFS and OS still remains .

ERG exploratory analyses results (II)

(deterministic results, PAS included)

Option	Total QALYs	Total Costs	Inc. QALYs	Inc. Costs	ICER
Exploratory analysis 4: standard care costs doubled (based on ERG-preferred model)					
Blinatumomab	7.35	[redacted]	2.75	£82,222	£29,848
Standard care	4.59	[redacted]	-	-	-
Exploratory analysis 5: alternative HSCT survival probabilities (based on ERG-preferred model)					
Blinatumomab	7.29	[redacted]	2.73	£89,302	£32,667
Standard care	4.55	[redacted]	-	-	-

Exploratory analysis 4: Alternative SoC costs: drug acquisition costs were doubled to assess the impact of assuming alternative treatment regimens. No significant impact on ICER

Exploratory analysis 5: Assess impact of alternative HSCT survival probabilities

- Shows that HSCT survival probabilities lead to an increase ICER for blinatumomab vs SoC; but ERG notes there is uncertainty around survival trajectory of HSCT patients

ERG exploratory analyses results (II) Alternative cure fractions for SoC and utilities

(deterministic results, PAS included)

Blinatumomab vs SOC	Inc. QALYs	Inc. Costs	ICER
Exploratory analysis 6 – alternative cure fractions for SoC (based on ERG-preferred model)			
Cure fraction = 0.21 (company base case)	2.75	£83,268	£30,227
Cure fraction = 0.25	2.36	£81,402	£34,465
Cure fraction = 0.30	1.83	£78,883	£43,072
Cure fraction = 0.35	1.30	£76,363	£58,697
Exploratory analysis 7 - Impact of alternative post-relapse utility values			
Utility = 0.69 (company's base case)	2.75	£83,268	£30,227
Utility = 0.819 (BLAST post-relapse utility)	2.67	£83,268	£31,157
Utility = 0.50	2.88	£83,268	£28,930
Utility = 0.25	3.04	£83,268	£27,395

ERG comments:

- Results show cure fraction is a key driver of cost-effectiveness for blinatumomab vs SoC
- Utility values for the post-relapse state have a minor impact on the ICER

ERG exploratory analyses – alternative models (III)

(deterministic results, PAS included)

Exploratory analysis 8 - Impact of using ERG's clinical advisors' preferred OS models		
OS model (low-high ICER determined by RFS curve)	Low ICER	High ICER
(a) Generalised gamma (unrestricted) preferred for blinatumomab arm by Clinical Advisor 1	£32,800	£34,904
(b) Restricted cubic spline Weibull (unrestricted) preferred for blinatumomab arm by Clinical Advisor 2	£30,868	£32,857
(c) Weibull mixture cure (unrestricted) selected for SoC by Clinical Advisor 1	£25,810	£27,492

ERG comment: Cure rate is driving cost-effectiveness. Inclusion of the 5-year cure assumption reduces variation in ICERs across the OS models considered (cure models also produce lower ICERs vs other OS forms)

- Range of low and high ICERs reflects the impact of assuming alternative RFS functions
- Only the Weibull non-mixture cure model (unrestricted) and the Weibull mixture cure model (unrestricted) produced ICERs below £30,000 per QALY gained

ERG Clinical experts: Cure point at 5 years is acceptable

- The distributions above were chosen based on:
  - (a) OS at 50% at 5 years data matched observed data from BLAST and MT103-202
  - (b) Provides clinically expected changes in OS between years 4 and 5
  - (c) The predicted 5-year OS probability
  - (d) RSC Weibull is preferred based on the fit to the (ATT-weighted) observed data up to five years.
- The clinical advisors' 3 preferred OS models result in ICERs in the range £25,810- £34,904 per QALY gained.

Innovation and equality

- Clinicians consider it innovative and a step-change in the management of ALL with MRD activity (Professional expert submission)
- Currently no targeted treatment option is available for people with MRD positive B-cell precursor positive ALL (Professional expert submission)
- Novel mechanism of action facilitates transient connection of malignant cells with T cells, thereby inducing T-cell-mediated killing of the bound malignant cell. By bringing T cells into close proximity with tumour cells much more frequently than without blinatumomab, the surveillance and cytotoxic abilities of the patient's own T cells are greatly increased (Company submission, B.2.12)
- No equality issues raised during scoping or company submission/ patient professional statements.

End of life criteria

Criterion	Data available
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	Median OS for the historical control group (using ATT-weighted propensity score matching analyses) for standard care chemotherapy was [redacted] The estimated mean survival (undiscounted) in the economic analysis was almost 5x greater than the median survival ([redacted] years) in the SoC arm; however, this is reflective of the small proportion of patients who achieve long-term survival (~20%).
There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment	Median OS (using ATT-weighted propensity score matching analyses), was [redacted] after more than 40 months follow-up for blinatumomab thus demonstrating a [redacted] OS survival [redacted] when compared to standard care. The estimated mean survival (undiscounted) in the economic analysis was [redacted] years in the blinatumomab arm, resulting in an incremental survival benefit of [redacted] years.
ERG comment	ERG disagrees with using median values to determine whether the end of life criteria are met. Medians represent the middle patient and don't take into account the skewness in the distribution of patient outcomes ERG's exploratory analyses show a lowest mean OS for the standard of care group of 7.69 years and a mean OS gain with blinatumomab of 2.12 years.

End of life considerations:
Landmark OS based on company's base case updated post-clarification

Landmark OS vs. BLAST				
Month	Blinatumomab		SOC	
	BLAST	Model	Historical Control	Model
6	[redacted]	[redacted]	[redacted]	[redacted]
12	[redacted]	[redacted]	[redacted]	[redacted]
24	[redacted]	[redacted]	[redacted]	[redacted]
53.5	[redacted]	[redacted]	[redacted]	[redacted]
60*		[redacted]	[redacted]	[redacted]
120*		[redacted]	[redacted]	[redacted]

**Input obtained from company model v0.4 by NICE technical team*

Key issues: cost effectiveness

1. Are cost-effectiveness results generalisable for the population in the MA, considering the absence of:

   (i) patients unable to receive HSCT or tolerate chemotherapy,

   (ii) patients in second complete remission (CR2)
2. Which parametric curves for OS and RFS are most appropriate for extrapolation?
3. How should cure be modelled? What cure point should be included in the model?

   Company preferred: no fixed cure point; ERG preferred: fixed cure at 5 years in both arms
4. Which post-relapse HRQoL estimate should be used: (i) observed utility of 0.692 among BLAST patients with post-relapse assessment, assumed values of (ii) 0.50 and (iii) 0.25
5. Does the model structure appropriately incorporate HSCT? Should alternative modelling be used or will it have similar uncertainty issues?
6. Which is the most plausible ICER?
7. End of life criteria
8. Equality and innovation
9. Suitable for CDF?

Authors

- **Lyudmila Marinova**
  Technical Lead
- **Alex Filby**
  Technical Adviser
- With input from the Lead Team: Stephen O'Brien, Alex Cale (clinical), Nigel Langford (cost), David Chandler (lay)

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Blinatumomab for acute lymphoblastic leukaemia for people with minimal residual disease activity in remission [ID1036]

Document B

Company evidence submission

October 2017

File name	Version	Contains confidential information	Date
ID1036_Blinatumomab MRD Company Submission Document B_October2017	1.0	Yes	31/10/2017

Contents

Tables and figures

Abbreviations

Abbreviation	Definition
ABL	Abelson murine leukaemia viral oncogene homolog
AC	appraisal committee
AE	adverse event
ALL	acute lymphoblastic leukaemia
ANC	absolute neutrophil count
AP	alkaline phosphatase
ASBMT	American Society for Blood and Marrow Transplantation
ASCO	American Society of Clinical Oncology
ASH	American Society of Hematology
ATE	average treatment effect
ATT	average treatment effect on the treated
BCP	B-cell precursor
BCR	B-cell receptor
BIC	Bayesian information criterion
BiTE	bispecific T-cell engager
BNF	British National Formulary
BSA	body surface area
CCR	Continuous haematological complete response
CDSR	Cochrane Database of Systematic Reviews
CENTRAL	Central Register of Controlled Trials
CHMP	Committee for Medicinal Products for Human Use
CI	confidence interval
CMV	cytomegalovirus
CNS	central nervous system
CR	complete response
CRF	case report form
CSF	cerebrospinal fluid
CSR	clinical study report
CTM	clinical trial material
CVAD	cyclophosphamide, vincristine, doxorubicin, dexamethasone
DARE	Database of Abstracts of Reviews of Effectiveness
DCAS	direct comparison analysis set
DNA	deoxyribonucleic acid
DSU	Decision Support Unit
EC	European Commission
ECCO	European Cancer Organisation
EFS	event-free survival

EHA	European Hematology Association
eMIT	Drugs and Pharmaceutical Electronic Market Information
EOI	events of interest
EORTC	European Organisation for Research and Treatment of Cancer
EPAR	European Public Assessment Report
EQ-5D	EuroQol 5 dimensions questionnaire
ESMO	European Society for Medical Oncology
EU	European Union
EWALL	European Working Group for Acute Lymphoblastic Leukaemia
FAS	full analysis set
FDA	Food and Drug Administration
FLAG-IDA	fludarabine, cytarabine, granulocyte colony-stimulating factor, idarubicin
GEE	generalised estimating equations
GLM	generalised linear model
GMALL	German Multicenter Acute Lymphoblastic Leukaemia Study Group
GOT	glutamic oxaloacetic transaminase
GPT	glutamic-pyruvic transaminase
HB	haemoglobin
HCHS	Hospital and Community Health Service
HCV	Hepatitis C Virus
HIV	Human Immunodeficiency Virus
HRG	healthcare resource group
HRQoL	health-related quality of life
HRU	healthcare resource utilisation
HSCT	haematopoietic stem cell transplantation
HTA	health technology assessment
ICER	incremental cost effectiveness ratio
IPTW	inverse probability of treatment weighting
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
IT	intrathecal
ITT	intention-to-treat
IV	intravenous
KM	Kaplan-Meier
LCL	lower confidence limit
LLOQ	lower limit of quantification
LYG	life years gained
MAA	marketing authorisation application
MRC	Medical Research Council
MRD	minimal residual disease
NA	not applicable
NCCN	National Comprehensive Cancer Network

NHS	National Health Service
NHS EED	NHS Economic Evaluation Database
NICE	National Institute for Health and Care Excellence
NMB	net monetary benefit
NR	not reported
NSAID	non-steroidal anti-inflammatory drug
OP	outpatient
OS	overall survival
PAS	Patient Access Scheme
PCR	polymerase chain reaction
PEG	polyethylene glycol
PPS	per protocol set
PR	post-relapse
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRO	patient-reported outcome
PRS	post-relapse survival
PSA	probabilistic sensitivity analyses
PSSRU	Personal Social Services Research Unit
QALY	quality-adjusted life year
QLQ	quality of life questionnaire
R	restricted
RCT	randomised controlled trial
RF	relapse-free
RFS	relapse-free survival
ROBINS	Risk Of Bias In Non-randomised Studies
RR	relative risk
RT	real-time
SD	standard deviation
SE	standard error
SLR	systematic literature review
SmPC	Summary of Product Characteristics
SoC	standard of care
STA	single technology appraisal
TCR	T-cell receptor
TKI	tyrosine kinase inhibitor
TTHR	time to haematological relapse
U	unrestricted
UCL	upper confidence limit
UK	United Kingdom
ULN	upper limit of normal
US	United States

WBC	white blood cells
WTP	willingness to pay

B.1 Decision problem, description of the technology and clinical care pathway

B.1.1 *Decision problem*

The submission covers the technology's full marketing authorisation for this indication.

Table 1 summarises the decision problem addressed in this submission.

Table 1: The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	People with B-cell precursor acute lymphoblastic leukaemia who have minimal residual disease (MRD) activity while in haematological remission	Adults with MRD+ B-precursor acute lymphoblastic leukaemia (ALL). Clinical evidence for blinatumomab is aligned with the proposed licensed indication; however, comparative evidence from a historical comparator study is limited to patients with Ph-negative B-precursor ALL who are in first complete haematological remission. Therefore, the economic analysis presented in this submission focused on this patient sub-group. Although the cost effectiveness evidence does not consider the Ph+ population or later remission states, due to the substantial unmet need across all sub-populations blinatumomab should be considered for use in alignment with its full anticipated marketing authorisation.	Blinatumomab is not expected to have a marketing authorisation for use in paediatric patients in this indication.
Intervention	Blinatumomab	Per final scope	NA
Comparator(s)	• Retreatment with combination chemotherapy • Monitor for relapse	• Retreatment with combination chemotherapy	Based on expert clinical opinion it is highly unlikely that MRD+ patients who have a high risk of relapse would solely be monitored for relapse without any

			treatment. Therefore, in the economic evaluation monitoring for relapse is not considered a comparator in its own right – instead, it is captured alongside ongoing chemotherapy regimens.
Outcomes	The outcome measures to be considered include: • Overall survival • Disease-free survival • Relapse-free survival • Minimal residual disease response • Rate of stem cell transplant • Adverse effects of treatment • Health-related quality of life	Per final scope	NA
Special considerations including issues related to equity or equality	If appropriate, the appraisal should include the costs associated with diagnostic testing for these cells in people with acute lymphoblastic leukaemia, while in remission, who would not otherwise have been tested. A sensitivity analysis should be provided without the cost of the diagnostic test. Guidance will only be issued in accordance with the marketing authorisation. Where the wording of the therapeutic indication does not include specific treatment combinations, guidance will be issued only in the context of the evidence that has underpinned the marketing authorisation granted by the regulator.	Per final scope	MRD status testing is already routine clinical practice in the diagnostic work-up and monitoring of BCP-ALL,[1, 2] and is recognised as an important marker for informing treatment decisions and prognosis. No additional tests or investigations are required for treatment with blinatumomab.

Abbreviations: NA: Not applicable
Source: NICE Blinatumomab Final Scope[3]

B.1.2 *Description of the technology being appraised*

A brief overview of blinatumomab is provided in Table 2.

Table 2: Technology being appraised

UK approved name and brand name	Blinatumomab (BLINCYTO®)
Mechanism of action	Blinatumomab is a first-in-class, bispecific T-cell engager (BiTE®) antibody construct that binds specifically to CD19 expressed on the surface of cells of B-lineage origin, and to CD3 expressed on the surface of T-cells. Blinatumomab activates endogenous T-cells by connecting CD3 expressed on the T-cell receptor complex with CD19 expressed on benign and malignant B-cells. Blinatumomab mediates the formation of a cytolytic immunological synapse between the T-cell and the malignant B-cell, triggering release of proteolytic enzymes to kill target cells. Blinatumomab is associated with transient upregulation of cell adhesion molecules, production of cytolytic proteins, release of inflammatory cytokines, and proliferation of T-cells, all of which results in elimination of CD19+ cells. It is the unique action of bringing T-cells into proximity with malignant B-cells much more frequently than without blinatumomab that greatly augments the surveillance and cytotoxic abilities of the patient's own T-cells. Thus, blinatumomab harnesses the body's own immune system to fight cancer.[4, 5]
Marketing authorisation/CE mark status	Blinatumomab was granted orphan designation by the European Commission (EC) in 2009.[6] A European marketing authorisation application (MAA) for blinatumomab in this indication was submitted in March 2017, and it is anticipated that the Committee for Medicinal Products for Human Use (CHMP) will adopt a positive opinion for this MAA in January 2018 for the indication of adults with MRD+ B-precursor ALL.
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	The anticipated indication for blinatumomab is for the treatment of adults with MRD+ B-precursor ALL. Blinatumomab also has an existing indication for the treatment of patients with Philadelphia chromosome-negative relapsed or refractory B-precursor ALL. For full details of the contraindications, warnings and precautions for use, see Appendix C.
Method of administration and dosage	Blinatumomab is administered by continuous intravenous (cIV) infusion delivered at a constant rate using an infusion pump. A single cycle of blinatumomab treatment comprises cIV infusion at a dose of 28 µg/day for 28 days, followed by a 14-day treatment-free interval. Step dosing is not required during the treatment of MRD+ BCP-ALL, unlike in the treatment of relapsed or refractory Ph-negative BCP-ALL, as the number of B-cells at baseline is low, as is the risk of cytokine release syndrome. Patients may receive 1 cycle of induction treatment followed by 3 additional cycles of consolidation treatment; treatment with blinatumomab should be discontinued if haematological relapse occurs. Based on an analysis of patients starting and completing each cycle in the pivotal phase II clinical trial, an average of 1.86 cycles of blinatumomab was received per treatment course, due in part to patients becoming eligible for and undergoing HSCT upon achievement of MRD-negativity (Section B.3.5).

Additional tests or investigations	MRD status testing is already routine clinical practice in the diagnostic work-up and monitoring of BCP-ALL,[1, 2] and no additional tests or investigations are anticipated to be required for treatment with blinatumomab.
List price and average cost of a course of treatment	The acquisition cost of blinatumomab is £2,017 per 38.5 µg vial (list price).[7] The average cost of blinatumomab per cycle at the list price is: • £56,476 (28 µg/day for Days 1–28, 28 vials)
Patient access scheme (if applicable)	A simple discount PAS has been approved by the Department of Health: • [redacted] • [redacted]

Abbreviations: BiTE: bispecific T-cell engager; CD: cluster of differentiation; EC: European Commission; MAA: marketing authorisation application; CHMP: Committee for Medicinal Products for Human Use; cIV: continuous intravenous; MRD: minimal residual disease; BCP: B-cell positive; ALL: acute lymphoblastic leukaemia; Ph: Philadelphia chromosome; PAS: Patient Access Scheme.
Source: Blinatumomab Summary of Product Characteristics, Nagorsen et al. (2009), Gökbuget et al. (2014), Hoelzer et al. (2016).[1, 4, 5, 8]

B.1.3 *Health condition and position of the technology in the treatment pathway*

Summary of Health Condition and Position of the Technology

- BCP-ALL is a rare form of leukaemia affecting relatively young adults; approximately 36% of patients will exhibit minimal residual disease (MRD) despite achieving a haematological complete response (CR) in the front-line setting (estimated n=85 in England and Wales).
- MRD is an independent predictive factor of outcome in BCP-ALL and is associated with high risk of relapse (both during first haematological CR and after salvage therapy) and poor survival outcomes.
    - In a recent meta-analysis by Berry et al. (2017) that included more than 13,000 patients, the predictive power of MRD status was confirmed; MRD+ patients were significantly less likely than MRD- patients to be disease-free (21% versus 64%), and alive (15% versus 60%) after 10 years.
- The primary goal of treatment is to achieve a cure, however, no approved treatments exist specifically for MRD+ BCP-ALL patients in haematological CR. Furthermore, currently available chemotherapy regimens are highly toxic and ineffective in achieving MRD negativity among BCP-ALL patients in haematological CR.
- HSCT, although associated with substantial morbidity and mortality, represents a potentially curative treatment option for high-risk patients, but importantly MRD+ patients experience significantly poorer outcomes.
- There is, therefore, a high unmet need for a targeted, effective treatment option that can realise a cure through achievement of MRD negativity among patients who are in haematological CR as well as improved post-HSCT outcomes, whilst reducing chemotherapy-associated toxicities in BCP-ALL patients.
- As the first and only drug indicated specifically for BCP-ALL patients, blinatumomab, by achieving MRD negativity, offers a paradigm shift in treatment options for these patients and is expected by expert clinicians to become the standard of care therapy in this setting: both prior to transplant in patients eligible for HSCT and as a stand-alone treatment for patients who are not eligible for HSCT.
- By eliminating MRD, which is an independent predictive factor for improved outcomes, patients receiving blinatumomab are substantially more likely to be cured of ALL, either via long-term, sustained molecular remission or receiving successful HSCT.

An overview of ALL and the position of blinatumomab in the current treatment pathway is provided in the following sections.

B.1.3.1 Disease overview

Acute lymphoblastic leukaemia

Leukaemia is a complex, progressive haematological malignancy that is characterised by the increased production of immature or abnormal blood cells by bone marrow and other blood-forming organs.[9] Acute lymphoblastic leukaemia (ALL) is a subset of leukaemia that refers to a group of haematopoietic neoplasms involving cells committed to the lymphoid lineage.[9]

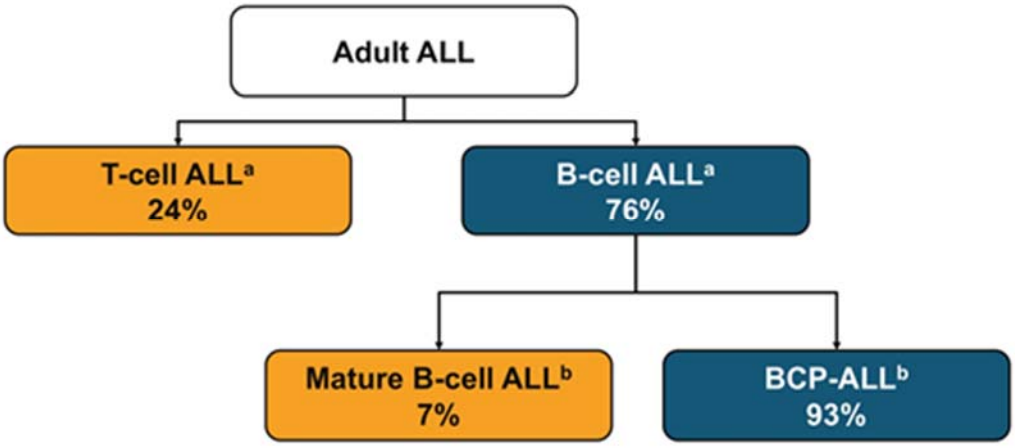
ALL specifically affects immature lymphocytes (lymphoblasts) that are derived from B- or T-lymphocyte stem cells. Proliferating lymphoblasts supress the production of normal blood cells in the bone marrow, causing haematological deficiencies including anaemia, immune system impairment, and platelet count deficiency.[10, 11] These leukaemic lymphoblasts express the same

antigens as normally developing B- and T-cells. Clinically, the recognised phenotypes are T-cells, mature B-cells, and precursor B-cells. Precursor B-cells typically express CD10, CD19, and CD34 cell surface markers.[10] Immunophenotyping is an important part of the diagnostic work-up for ALL to classify cases, for immunologic monitoring of MRD (defined as the detection of more than 1 cancerous cell per 10,000 normal cells), and for treatment with targeted cellular immunotherapy.[10]

ALL sub-classifications

Although the aetiology of ALL is unclear, it is one of the most carefully studied and best-characterised neoplasms. An overview of the ALL sub-classifications is provided in Figure 1; the majority of ALL cases (76%) are B-cell lineage, of which 93% are B-cell precursor (BCP) ALL.[12-14] BCP-ALL is the population from which patients relevant to this appraisal are drawn.[3] Philadelphia chromosome-positive (Ph+) is a biologically and clinically distinct variant of ALL classified as ALL with translocation t(9;22)(q34;q11.2) and accounts for 20% to 30% of ALL cases in adults and 2% to 3% of ALL cases in children (across all sub-classifications, not specifically BCP).[13, 15]

Figure 1. ALL sub-classifications



Footnotes: [a]Percentages were derived by calculating weighted averages of the proportion of adult ALL that is B-cell lineage.[16-20] [b]Percentages were derived by calculating weighted averages of the proportion of adult B-lineage ALL that is BCP-ALL.[13, 14]
Abbreviations: BCP-ALL: B-cell precursor ALL
Source: Amgen ALL Epidemiology Estimates (Appendix N)

MRD in current clinical practice

MRD refers to residual ALL present at frequencies below the sensitivity of standard microscopy, but detectable by molecular means such as PCR or flow cytometry, in the bone marrow of patients who have met the criteria for haematological CR.[21] In adult ALL, more than 80% of Ph-negative patients respond to induction chemotherapy with a haematological CR, and yet 44% of these patients experience relapse at a median of 11 months from the start of treatment.[13] Thus, despite an impressive 40% to 50% overall survival rate at 5 years, a prognosis achieved over the last 10 years, refractory relapsed leukaemia remains an unsolved therapeutic problem.[13] Among patients with ALL in their first CR, prognostic factors for relapse include baseline features such as cytogenetics (particularly the 9;22 translocation), white blood cell (WBC) count, and age.[17, 22, 23]. However, the persistence of MRD has been shown to be the strongest predictive factor for

relapse regardless of treatment choice or risk classification system;[2, 24-26] as presented in Table 3, patients in one study who failed to achieve MRD- status experienced a greater 3-year relapse rate, while those who achieved MRD clearance rapidly experienced a 3-year relapse rate of 0%.[26] At the time of relapse, the strongest prognostic factors for overall survival (OS) are duration of initial haematological remission and age.[13, 18, 27]

Table 3. The effect of MRD status on 3-year relapse rate in Brüggemann et al. (2006)[26]

MRD Risk Group	3-Year Relapse Rate (95% CI)
Low risk (10% of patients)	0% (NA)
Intermediate risk (67% of patients)	47% (31% – 63%)
High risk (23% of patients)	94% (83% – 100%)

Footnotes: Low risk was defined as a rapid MRD decline to lower than 10^{-4} or below detection limit at day 11 and day 24, high risk was defined as an MRD of 10^{-4} or higher until week 16, the remaining patients were defined as intermediate risk.
Abbreviations: MRD: minimal residual disease; CI: confidence interval; NA: not applicable.
Source: Brüggemann et al. (2006)[26]

The presence of MRD is a continuous variable, and patients who are highly responsive to induction chemotherapy and who achieve an MRD level below 1 x 10^{-4} (MRD- based on the sensitivity of the methodology) have a favourable prognosis. Typically, the presence of MRD represents disease that is insensitive to the multi-agent therapy used for induction and/or consolidation chemotherapies (see Section B.1.3.2), and thus subsequent rounds of similar therapy may not be efficacious for eliminating MRD.[21]

Assessment of MRD is commonly used clinically to evaluate the depth of response, categorise the level of risk of relapse, and to aid in treatment decisions.[28] MRD is evaluated by multiple methods, most commonly multichannel flow cytometry with immunophenotypic markers, real-time quantitative polymerase chain reaction (RT-qPCR), or next-generation sequencing. Patients are considered to have MRD (*i.e.*, to be MRD+) if molecular evidence of blasts in the bone marrow is detectable above the lower limit of quantitation of 1 x 10^{-4} (> 1 in 10,000). An MRD level > 1 x 10^{-4} is deemed MRD+ and represents a very high-risk condition for relapse.[29] The method and timing of MRD testing varies, and is described in more detail in Section B.1.3.3.

Incidence and prevalence of the MRD population

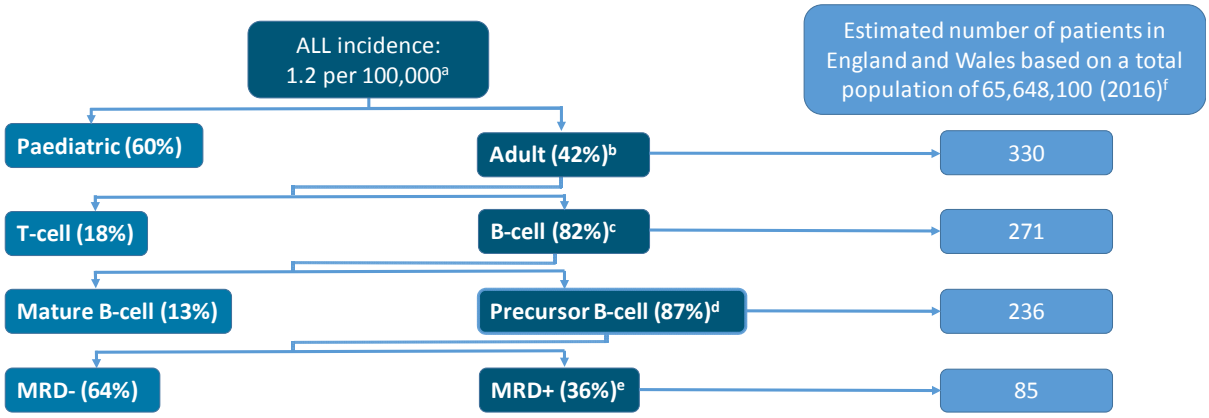
ALL is a rare disease, with an incidence in the UK of approximately 1.2 per 100,000, with 758 new cases diagnosed across the UK in 2014.[30] The incidence has its peak during childhood, decreasing with increasing age. From the age of 35 years on the incidence rises again and a second peak is observed starting from the age of 80 years.[30]

The proportion of patients in MRD+ haematological CR after front-line chemotherapy can vary between studies due to the timing of MRD testing (after induction or consolidation, see Section B.1.3.3). The use of different MRD testing methodologies and MRD thresholds can also lead to variability. Nonetheless, a weighted analysis of large prospective, multicentre studies have reported that 33–47% of patients have MRD after induction therapy;[19, 25, 31] a weighted analysis of these trials has produced a rate of 36%.[12]

Figure 2 shows the estimated number of cases of ALL in the UK, according to sub-population, calculated by applying the estimated incidence of ALL to UK population estimates. In an estimated population of 236 patients with BCP-ALL, there are expected to be 85 patients who are MRD+ after receiving front-line chemotherapy.[12] These patients would have a higher risk of

relapse and shorter survival than MRD- patients. Effective treatment that achieves MRD negativity in patients achieving haematological CR after front-line chemotherapy would reduce the number of patients who experience a relapse.

Figure 2. Estimated incidence of adult MRD+ B-precursor- ALL in England and Wales



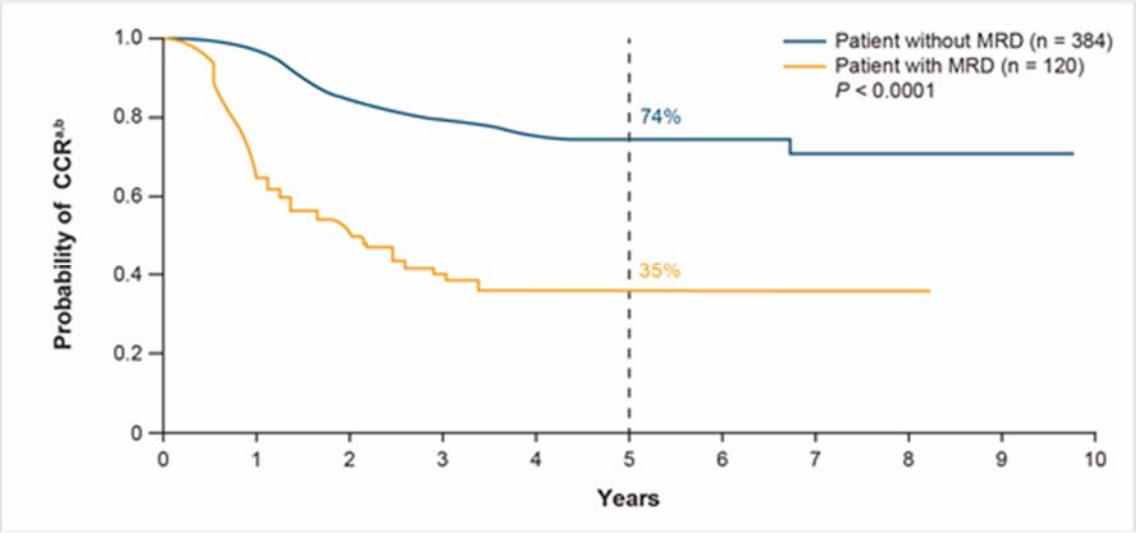
Footnotes: [a]Cancer Research UK (2016 estimate)[30]
[b]Calculated from UK age-specific ALL incidence data reported by Cancer Research UK (2011-2013 estimate).[30] Since data were only provided for 5-year age groups, the 15-19 year age group was split such that 60% of the population projection for this age group was considered 15-17, and the remaining 40% were considered 18-19 and included in the estimate for adult patients.
[c]Weighted average of data from (i) a UK cytogenetic population-based study of 349 patients (> 15 years of age) with ALL diagnosed between 1983 and 2001 (Moorman et al., 2010);[14] and (ii) an analysis of cytogenetic data from 1522 patients (15 years to 65 years of age) with ALL enrolled on the MRC UKALLXII/ECOG 2993 study (Moorman et al., 2007).[17] Data on T- and B-cell lineage from Moorman et al., 2007 was calculated using separately reported proportions of patients with T-cell lineage in subsets of patients with Ph+ ALL and Ph- ALL.
[d]Based on UK data from a cytogenetic population-based study of 349 patients (>15 years of age) with ALL diagnosed between 1983 and 2001 (Moorman et al., 2010).[14]
[e]Amgen ALL epidemiology estimates (Appendix N)

Prognosis and unmet need for MRD+ patients

The prognosis for patients treated with currently available therapies is dependent on a number of factors: most long-term survivors of ALL have undergone HSCT and have other well-established positive prognostic factors, including a younger age, shorter time to CR, longer duration of CR, later relapse, and lower white blood cell counts.[18, 27, 28, 32, 33]

MRD is increasingly considered an independent predictive marker for duration of response and long-term outcomes in patients with ALL, and is important for assessing the risk of relapse and informing treatment decisions. Adult MRD+ BCP-ALL patients in haematological CR have an increased risk of haematological relapse and death compared with those who do not have MRD. In a large German Multicenter ALL study (GMALL), the probability of maintaining haematological CR without relapse at 5 years was 35% for MRD+ patients after front-line chemotherapy, compared with 74% of MRD- patients ($p < 0.001$), as shown in Figure 3.[19]

Figure 3. Continuous Haematological CR (CCR) without relapse



Footnotes: [a]The probability of CCR was calculated from the date of achieving haematological complete remission to the date of relapse or last follow-up. [b]Includes patients who underwent HSCT.
Abbreviations: CCR: continuous haematological complete remission; HSCT: haematopoietic stem cell transplant
Source: Gökbuget et al. (2012)[19]

In the same study, MRD+ patients also had significantly poorer OS than MRD- patients, with a 5-year OS of 42% versus 80% (p = 0.0001).[19] In a recent meta-analysis by Berry et al. (2017) that included more than 13,000 patients, the predictive power of MRD status was confirmed; MRD+ patients were significantly less likely than MRD- patients to be disease-free (21% versus 64%), and alive (15% versus 60%) after 10 years. Furthermore, the predictive power of MRD status is substantial and robust, irrespective of ALL sub-population or MRD detection method, period or cut-off level, as shown in Table 4.

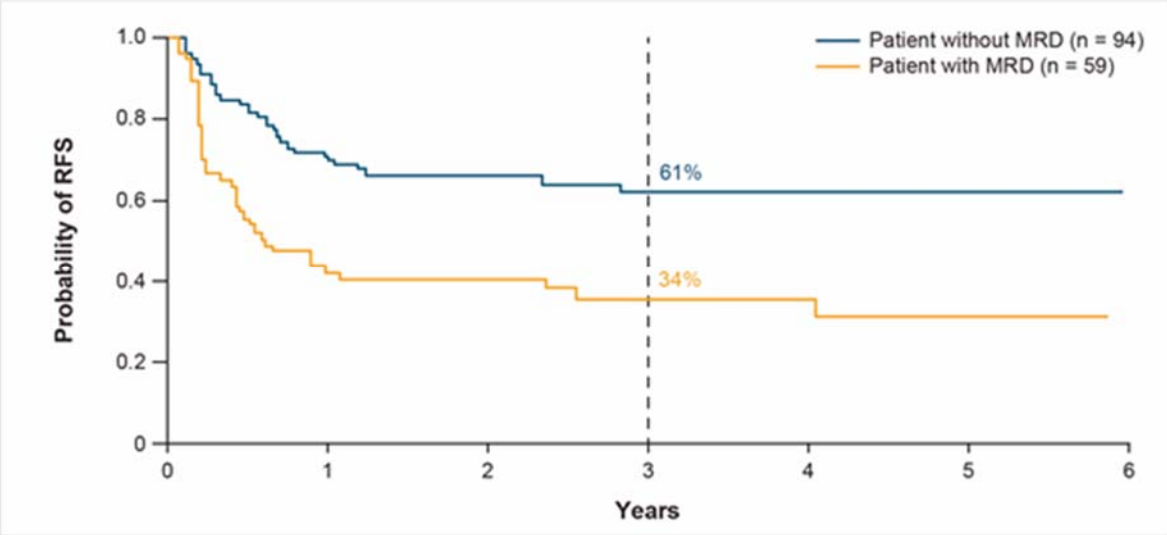
Table 4: Predictive effect of MRD status on EFS and OS

Subgroup	EFS, Hazard Ratio (95% CI)	OS, Hazard Ratio (95% CI)
MRD Detection Method		
Flow cytometry	0.32 (0.20, 0.51)	0.28 (0.13, 0.61)
PCR	0.24 (0.18, 0.32)	0.29 (0.18, 0.49)
MRD Cut-off		
0.0001	0.29 (0.21, 0.39)	0.25 (0.18, 0.35)
<0.0001	0.21 (0.14, 0.32)	0.30 (0.18, 0.50)
MRD Detection Period		
Induction	0.33 (0.24, 0.44)	0.54 (0.24, 1.20)
Consolidation	0.25 (0.18, 0.36)	0.27 (0.18, 0.40)
Other period	0.18 (0.08, 0.41)	0.20 (0.04, 0.92)
Cytogenetics		
Ph-	0.28 (0.22, 0.37)	0.26 (0.17, 0.40)
Ph+	0.34 (0.22, 0.53)	0.38 (0.19, 0.75)
Cell Phenotype		
B-cell	0.28 (0.17, 0.45)	NR
T-cell	0.31 (0.19, 0.53)	NR

Abbreviations: MRD: minimal residual disease; EFS: event-free survival; OS: overall survival; CI: confidence interval; PCR: polymerase chain reaction; Ph: Philadelphia chromosome.
Source: Berry et al. (2017)[34]

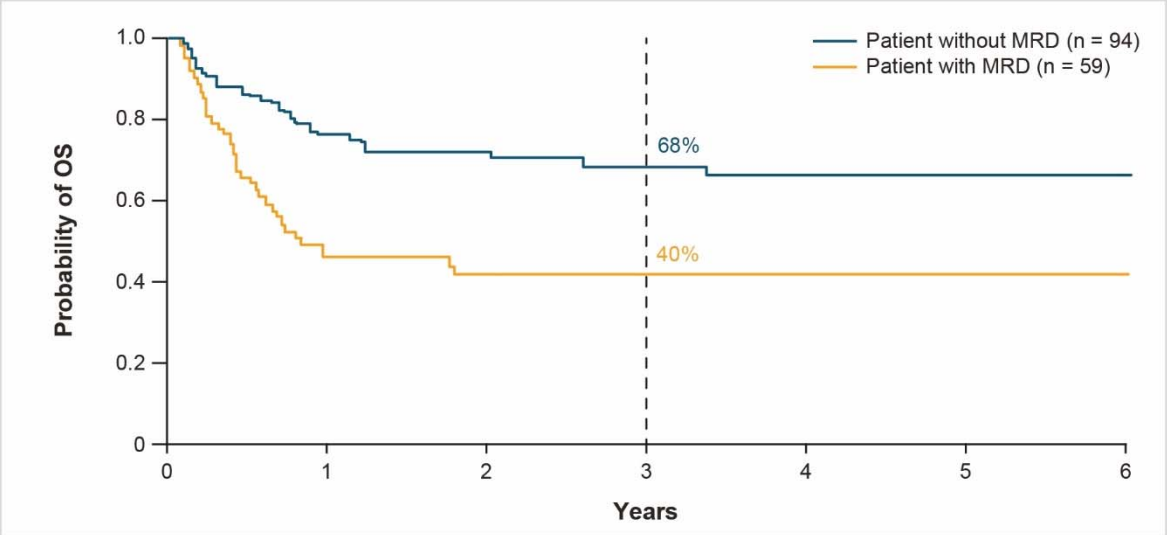
HSCT is a potentially curative treatment option for high-risk patients after induction therapy. However, as described in more detail Section B.1.3.2, it is associated with severe morbidity and a high mortality rate, and the risk of HSCT failure (i.e. haematological relapse) is greater in MRD+ BCP-ALL patients. As shown in Figure 4 and Figure 5, MRD+ patients have a substantially higher risk of haematological relapse and death at 3 years' post-transplant, compared with MRD- patients.[35]

Figure 4. RFS after HSCT in MRD+ and MRD- patients



Footnotes: Of 153 patients, 142 (89%) received total body irradiation-based conditioning; the remainder received regimens consisting of treosulfan and fludarabine, busulfan and cyclophosphamide, or busulfan and fludarabine, and then underwent HSCT.
Abbreviations: HSCT: haematopoietic stem cell transplant; OS: overall survival; RFS: relapse-free survival.
Source: Bar *et al.* (2014)[35]

Figure 5. OS after HSCT in MRD+ and MRD- patients



Footnotes: Of 153 patients, 142 (89%) received total body irradiation-based conditioning; the remainder received regimens consisting of treosulfan and fludarabine, busulfan and cyclophosphamide, or busulfan and fludarabine, and then underwent HSCT.
Abbreviations: HSCT: haematopoietic stem cell transplant; OS: overall survival; RFS: relapse-free survival.
Source: Bar *et al.* (2014)[35]

MRD- status prior to HSCT is associated with stronger post-transplant outcomes, improving the risk/benefit ratio. Despite the poorer post-transplant outcomes for patients who have MRD+,

HSCT is still considered to be the best available treatment option for these patients due to the substantial risk of relapse faced by this population and because HSCT is the only treatment to be potentially curative.[1, 28] However, MRD+ patients are at a high risk of relapse and treatments that delay or postpone relapse could facilitate an increase in the number of transplants, as fewer patients would relapse before identifying suitable donors. Furthermore, treatments that achieve MRD negativity and demonstrably sustain this response over time could conceivably be considered a suitable alternative to transplant in clinical practice in future.

Ultimately, BCP-ALL patients who do experience a haematological relapse not only face imminent risk of death, but also a substantial reduction in HRQoL; in the large Phase III TOWER trial in adults with relapsed or refractory BCP-ALL, more than 50% of patients reported being more than moderately tired and physically weak at baseline; at least 25% of patients reported their tiredness and physical weakness to be greater than 'quite a bit' at baseline, and at least 25% of patients had joint or bone pain, were unable to eat, and had night sweats sometimes over the last 7 days.[36]

Furthermore, the economic burden of BCP-ALL is high due to the increased healthcare resource use, including lengthy and repeated hospitalisations; in a French study, during the period of salvage chemotherapy adults with relapsed or refractory BCP-ALL spend almost half of their time (46%) in hospital, resulting in high healthcare resource use and costs, and increasing the burden on the healthcare system.[37] Similar studies in Belgium, Germany, Italy, Spain, and the US found that adults with relapsed or refractory BCP-ALL spent more than half their time in hospital, placing a significant economic burden on the healthcare system.[38-42]

A summary of the burden of MRD in BCP-ALL is presented in Figure 6.

Figure 6. Burden of MRD in BCP-ALL

Footnotes: [a]Compared with adult BCP-ALL patients in haematological CR and MRD-. Schematic is based on published evidence.[5, 19, 37-43]
Abbreviations: BCP-ALL: B-cell precursor ALL; HRQoL: health-related quality of life; HSCT: haematopoietic stem cell transplant.

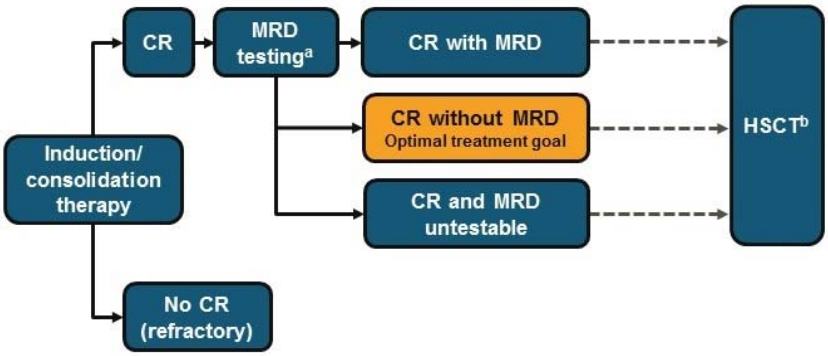
Achievement of MRD- status correlates positively with CR duration, reduced risk of relapse, and increased success of HSCT (and therefore the chance of achieving a cure for ALL);[25, 26] therefore, reducing and maintaining MRD below the lower limit of quantification (LLOQ) is the optimal treatment goal for MRD+ BCP-ALL patients in haematological CR.

B.1.3.2 Treatment aims and current treatment options

The primary treatment goal for BCP-ALL is to achieve a cure through sustained MRD negativity (below the lower limit of quantification, i.e. 10^{-4}) and maintained haematological CR, which is

defined as a bone marrow blast level of < 5% (*i.e.* undetectable by light microscopy);[1] currently most long-term survivors achieve such a cure by undergoing HSCT. Given the high relapse rate among MRD+ BCP-ALL patients, an MRD- haematological CR (i.e., molecular CR) status is the optimal outcome, as shown in Figure 7.[1] Currently, no approved treatments exist specifically for MRD+ BCP-ALL patients in haematological CR.

Figure 7. Treatment goal for ALL



Footnotes: [a]Timing of MRD assessment varies depending on treatment protocol used and can include end of induction and post-induction phase; [b]The decision to send eligible patients to HSCT is dependent on several factors, including presence of other high-risk factors; the role of HSCT in high-risk patients in haematological CR without MRD is not confirmed.[1]
Abbreviations: CR: haematological complete remission; HSCT: haematopoietic stem cell transplant; MRD: minimal residual disease.

Currently, treatments used to induce haematological remission in adult patients with Ph-negative ALL are typically comprised of blocks of multi-agent therapy regimens with different combinations or variations of cytotoxic, antineoplastic, and other agents.[28, 44]. In addition, intrathecal chemotherapy, with or without radiation to the brain, forms part of the treatment regimen to treat or prevent central nervous system (CNS) relapse. Ph-positive ALL is typically treated with similar agents with the addition of a tyrosine kinase inhibitor (TKI). The choice of initial chemotherapeutic agents depends on several factors, including the adverse event profile of therapeutic options, patient comorbidities, performance status, regional practice pattern, and physician preference. In the UK, treatment of newly diagnosed ALL is based primarily on the UKALL14 trial,[45] as described in more detail in Section B.1.3.3.

Although more than 80% of patients achieve haematological CR after induction therapy, up to 44% of patients will relapse at a median of 11 months from the start of treatment,[13] as described in Section B.1.3.1.[29, 46] Furthermore, chemotherapies used in the treatment of ALL are highly toxic and patients who relapse may not be able to tolerate a new round of intensive salvage chemotherapy. Rates of adverse events (AEs) with chemotherapies to treat relapsed ALL are high: a systematic review of AEs in trials of chemotherapy regimens for relapsed or refractory BCP-ALL found that almost all patients treated with standard combination chemotherapy regimens experienced haematological toxicity (cytopaenia, neutropaenia, thrombocytopaenia) and rates of grade ≥3 haematological toxicity were high for most treatments. Infections were a common toxicity; mucositis and gastrointestinal toxicities were also common.[47]

MRD status is a major predictive factor for relapse among patients in haematological CR, and as described in Section B.1.3.1, a weighted analysis of prospective, multicentre trials determined that 36% of patients who achieve haematological CR maintain MRD+ status after front-line

chemotherapy.[12] Therefore, the majority of study groups in Europe recommend allogeneic HSCT for eligible patients with MRD+ ALL after consolidation treatment, as this is the most intensive and potentially curative treatment option.[1] Unfortunately, the outcome of allogeneic HSCT in MRD+ patients is suboptimal; in a study of patients who underwent allogeneic HSCT, patients who were MRD+ prior to HSCT demonstrated a 36-month OS of 49%, compared to 80% in patients who were MRD- prior to HSCT.[48] Similarly, the cumulative incidence of post-HSCT relapse for patients who are MRD+ prior to HSCT is 46% compared to 0% for patients who are MRD- prior to HSCT.[48] Therefore, achieving MRD- status in these patients could be expected to improve patient survival. In addition, the use of HSCT as a treatment option in BCP-ALL is variable across clinical practice due to clinician and patient preferences, particularly in light of the risk associated with undergoing transplantation and the poor post-transplant outcomes associated with HSCT in patients who are MRD+. Furthermore, not all patients are suitable candidates for HSCT, due to, for example, age, medical comorbidities, or lack of a suitably-matched donor.[1] Therefore, for those MRD+ patients not eligible for HSCT, current guidelines recommend the continuation of first-line chemotherapies, which are ineffective in eliminating MRD, thereby leaving patients at high risk of relapse, and which are associated with a substantial adverse event profile.[25, 46]

The prognosis for MRD+ BCP-ALL patients in haematological CR is extremely poor, yet current chemotherapy options are highly toxic and ineffective at achieving MRD negativity.[26, 29, 46] Furthermore, there are no treatment options specifically indicated for MRD+ BCP-ALL aside from HSCT, which itself is associated with substantial morbidity and mortality.[49,50] The presence of MRD in a patient with haematological CR is recognised as the most important predictive factor for relapse and death.[48] While allogeneic HSCT is an option, some patients are ineligible (due to age, comorbidities or lack of donor), and MRD+ patients have a significantly higher risk of relapse post-transplant than MRD- patients.[48]

Therefore, MRD+ patients have limited treatment options to prevent haematological relapse, and as such possess an extremely poor prognosis and substantial increase in the risk of death.[50]

There is, therefore, a high unmet need for a targeted, effective treatment option that can achieve MRD negativity and sustain haematological CR, in addition to improving post-HSCT outcomes, and reducing chemotherapy-associated toxicities in BCP-ALL patients.

B.1.3.3 Clinical guidelines and treatment pathway

Clinical Guidelines

There are currently no published NICE clinical guidelines relevant to the management of adult MRD+ BCP-ALL.

Pegaspargase was recommended in NICE TA408 as a treatment option for children, young people, and adults with ALL, but the manufacturer submission and subsequent recommendation was limited to patients with newly-diagnosed disease.[51] This technology appraisal is therefore not considered relevant to the current appraisal.

Two other potentially-relevant NICE TAs are planned/in development:

- 'Erythrocyte encapsulated asparaginase for treating acute lymphoblastic leukaemia in adults and children after treatment with *Escherichia coli*-derived asparaginase' (for the treatment of

people with ALL who are intolerant or allergic to asparaginase, or have disease that has relapsed on asparaginase treatment) [ID864]. Suspended as of 28 October 2016

- 'Inotuzumab ozogamicin for treating relapsed or refractory acute lymphoblastic leukaemia' [ID893].

The European Society for Medical Oncology (ESMO) guidelines for ALL describe the importance of MRD testing, and the achievement of MRD- status as the most relevant predictive factor for disease-free survival and OS. The guidelines recommend MRD testing at the end of induction therapy and in the post-induction phase to evaluate treatment response, and every 3 months in the follow-up of asymptomatic patients. However, these guidelines do not suggest treatment options for patients in MRD+ haematological CR beyond HSCT and standard chemotherapy regimens.[1]

Treatment Pathway

As described previously, treatment of ALL in the UK is typically based on the UKALL14 protocol.[45] Treatment begins with induction therapy, the primary goal of which is the complete eradication of ALL cells from the blood, bone marrow and CNS or other extramedullary sites (when initially involved). This should be achieved as rapidly as possible in order to start post-haematological remission consolidation therapy. For Ph-negative ALL, induction therapy involves three sequential, connected steps: a pre-phase, induction I and induction II, with the latter applied regardless of CR after induction I.

Newly diagnosed BCP-ALL patients are first treated with a steroid pre-phase of 5–7 days, followed by induction I, in which all patients (regardless of phenotype) receive daunorubicin, vincristine, dexamethasone, PEG-asparaginase and methotrexate. Treatment then progresses to induction II, with all patients receiving cyclophosphamide, cytarabine, mercaptopurine and methotrexate. Patients achieving haematological CR at this stage may then progress to HSCT (if considered high risk, e.g. MRD+, clinically eligible, willing to undergo HSCT and a suitable donor is available), with or without intensification. Intensification therapy consists of methotrexate and PEG-asparaginase. Consolidation therapy is given to patients not eligible for HSCT, with 4 cycles variously utilising cytarabine, etoposide, PEG-asparaginase, methotrexate, daunorubicin, vincristine, dexamethasone, cyclophosphamide, cytarabine and mercaptopurine. Finally, maintenance therapy consisting of vincristine, prednisolone, mercaptopurine and methotrexate is given for 2 full years. Ph+ patients additionally receive daily imatinib throughout the induction, intensification, consolidation and maintenance phases of treatment.[45]

MRD testing

As described above, treatment guidelines show that MRD testing is an essential part of the patient management process for ALL.[1, 28] A recent survey of physicians in the UK also affirms the importance and widespread implementation of MRD testing in current NHS practice.[52] MRD can be evaluated in approximately 95% of patients with ALL.[53] Although there is not a universally established measure of MRD, it is commonly defined as the presence of 0.01% (10^{-4}) or more ALL cells in the bone marrow.[21, 28] However, clinical studies have defined MRD using various thresholds and time points.

Recommendations on the minimum technical requirements for assessing MRD were developed by a consensus development workshop in 2008.[54] A number of different technologies are

available to measure MRD, with varying specificity and sensitivity.[2, 53, 55] The most common methods include:[54]

- Multicolour flow cytometry to detect abnormal immunophenotypes, with a sensitivity of 10^{-3} to 10^{-4} for 3 to 4 colour flow cytometry and 10^{-4} to 10^{-5} for 6 to 9 colour flow cytometry
- Real-time quantitative polymerase chain reaction (RT-qPCR) assays to detect clonal rearrangements in Ig heavy chain genes, and/or TCR genes, with a sensitivity of 10^{-4} to 10^{-5}
- RT-qPCR assays to detect fusion genes (e.g., BCR–ABL), with a sensitivity of 10^{-4} to 10^{-6}

Although treatment guidelines for ALL recommend MRD testing to be routinely conducted in patients in haematological CR, global consensus has not yet been reached on precisely when to test for MRD. The US National Comprehensive Cancer Network (NCCN) guidelines recommend that patients who experience a haematological CR after induction therapy should be monitored for MRD on completion of initial induction and at additional time points depending on the regimen used.[28] Similarly, ESMO guidelines recommend MRD testing at the end of induction therapy and in the post-induction phase to evaluate treatment response, and every 3 months in the follow-up of asymptomatic patients.[1]

There is, however, an apparent consensus on MRD testing patterns in the NHS, as found in a recent survey of clinical practice in the UK, which confirmed the importance and prevalence of MRD testing of BCP-ALL patients. A group of 20 physicians were recruited, the majority of whom participated in research (75%), and who usually treated their patients according to the UKALL14 research protocol (70%). The survey found that among patients who achieved CR with front-line treatment (CR1), MRD testing was conducted in 70% of their patients. For patients who achieved CR2 or later, MRD testing was reported in 58% of patients.[52]

An initial prognostic MRD test most commonly commenced 4–8 weeks after the start of induction therapy (79%). Following the prognostic MRD test, the median estimate for the number of post-CR MRD tests in an individual patient over the subsequent 12 months was four for patients who were MRD- with Ph- disease, or MRD- or MRD+ with Ph+ disease, and three for patients who were MRD+ with Ph- disease. The average reported frequency of testing was every 3 months for patients who were MRD-, irrespective of Ph status, every 5 months for patients who were MRD+ with Ph- disease, and every 4 months for patients who were MRD+ with Ph+ disease. The majority of physicians (56–67% depending on Ph and MRD status) stated that the number of tests aligned with the protocol they followed.[52]

These findings demonstrate current widespread use of MRD testing throughout treatment of BCP-ALL in the NHS. Overall, there is value in the measurement of MRD status as early as possible and over multiple timepoints through the treatment process; Brüggemann et al. (2006) determined that identifying MRD status at different times resulted in differing outcomes, with those achieving MRD negativity during induction experiencing improved RFS and OS than those achieving MRD negativity after induction.[26]

While international and local guidelines differ on the timing of MRD testing, the available clinical evidence demonstrates the importance of determining MRD status and achieving MRD negativity; in their meta-analysis of more than 13,000 patients, Berry and colleagues (2017) demonstrated significantly better event-free survival and OS in patients who achieve MRD-status, irrespective of disease sub-type or MRD testing method.[34] Therefore, clinical practice

should aim to identify MRD status as early as possible in the treatment pathway and target treatment towards achieving MRD negativity as early as possible.

B.1.3.4 Proposed use and positioning of blinatumomab

Blinatumomab is the first and only drug indicated specifically for MRD+ BCP-ALL patients in haematological CR. The marketing authorisation for blinatumomab is anticipated to encompass both the Ph-negative and Ph-positive populations; however, due to the orphan nature of MRD+ BCP-ALL, comparative efficacy data are only available in the Ph-negative population in first CR (CR1), which is a considerably larger sub-population than the Ph-positive group (<5% patients Ph+ in BLAST). As such, while the cost effectiveness evidence presented in Section B.3 considers only the Ph-negative CR1 population, due to the substantial unmet need across both Ph-negative and Ph-positive sub-populations, blinatumomab should be considered for use in alignment with its full anticipated marketing authorisation.

The comparator included in this submission, retreatment with combination chemotherapy, is aligned with the final NICE scope (see Section B.1.1), and with UK clinical practice (the UKALL14 protocol, as described in Section B.1.3.3). Consequently, HSCT is not considered a comparator in this submission, implicitly assuming that blinatumomab will not displace HSCT and is instead likely to be used prior to HSCT in patients eligible to undergo transplant or to delay the need for HSCT. Nonetheless, by achieving and sustaining MRD negativity over time, blinatumomab may conceivably delay transplant indefinitely in clinical practice in future.

Blinatumomab is currently licensed and NICE-approved in the relapse/refractory setting and will be the first drug indicated specifically for BCP-ALL patients in MRD+ haematological CR. Given the high unmet need in this rare condition, UK clinical expert opinion consistently supports the use of blinatumomab as early as possible in the treatment pathway, with initiation after front-line chemotherapy (*i.e.* after 2 induction cycles) considered to be the most appropriate timepoint. In clinical practice, blinatumomab is expected to displace continued chemotherapy regimens and/or be used prior to transplantation depending on the most appropriate treatment pathway for the patient. Subsequent use of blinatumomab to treat MRD positivity in later remission states or as a salvage therapy is not anticipated if blinatumomab is used in the aforementioned setting.

A major benefit of blinatumomab is the ability to achieve sustained MRD response, which is associated with an improved prognosis independent of transplant, and a reduction in toxicities associated with conventional chemotherapy regimens; in addition, blinatumomab improves prognosis in patients post-HSCT. By substantially improving survival and post-HSCT outcomes, blinatumomab is expected to substantially increase the number of patients being cured of ALL.

As no other treatment options are specifically indicated for MRD+ BCP-ALL patients in haematological CR beyond continued chemotherapy, blinatumomab offers hope to patients and a targeted and effective therapeutic option to prescribers; blinatumomab is not just another incremental expansion of the therapeutic armamentarium for ALL but represents a significant paradigm shift in treatment for this rare and deadly disease. Treatment with blinatumomab eradicates MRD in a high proportion of patients, resulting in improved and sustained RFS and OS, and substantially increases the likelihood for patients to achieve a potential cure. Blinatumomab is, therefore, expected by clinicians consulted by Amgen to become the SoC for this population.

B.1.4 *Equality considerations*

No equality issues relate to the use of blinatumomab for the treatment of adult MRD+ BCP-ALL patients in haematological CR.

B.2 Clinical effectiveness

Summary of Clinical Effectiveness

- A systematic literature review was conducted, and 3 relevant studies were captured:
  - Two single-arm studies (MT103-202, BLAST) and one historical comparator study (Study 20120148)
- The pivotal single-arm trial, BLAST, demonstrated blinatumomab to reduce MRD below the LLOQ (usually 10^{-4}, as shown in Table 10) in 78% of MRD+ BCP-ALL patients in haematological CR within the first cycle.
- Patients who achieved complete MRD response with blinatumomab had 17.9 months longer median haematological RFS and a median OS more than triple that of patients who did not.
- Patients treated with blinatumomab in first haematological remission had more than double the median haematological RFS of those in later remissions.
- Blinatumomab offers durable complete MRD responses, with a median duration of 17.3 months.
- The highly effective results of blinatumomab treatment were also reflected in the pilot Phase II trial (MT103-202; n=20):
  - 80% of evaluable patients achieved MRD response, with all MRD responses having been observed within the first treatment cycle.
  - Median duration of complete MRD response for patients was 13.0 months.
  - Median haematological RFS had not been reached after a median follow-up time of 1550 days (> 4 years).
  - After up to 5.9 years follow-up 52.6% of patients treated with blinatumomab remained relapse-free.
- Due to the single-arm nature of BLAST, a historical comparator study was performed to provide a selected and well-matched cohort of patients treated with SoC, therefore permitting comparison to blinatumomab using propensity score analysis.
- Compared to a historical cohort treated with SoC, blinatumomab reduces the risk of haematological relapse or death by [redacted] , and more than quintuples the median RFS in MRD+ BCP-ALL patients in haematological CR.
- Blinatumomab reduces the risk of death by [redacted], and [redacted] the median OS in MRD+ BCP-ALL patients in haematological CR.
- A higher proportion of blinatumomab-treated patients [redacted] underwent, HSCT than historical control patients receiving SoC chemotherapy [redacted]
- Results were similar when patients were censored at HSCT, demonstrating that achievement of MRD negativity is an independent predictive factor for positive outcomes.
- No new safety signals were observed in patients pooled from BLAST and MT103-202, beyond the existing safety profile of blinatumomab.
- Neurological events occurred in patients receiving blinatumomab, however, more than two thirds of the neurological adverse events associated with blinatumomab were mild to moderate and decreased over time.

B.2.1 *Identification and selection of relevant studies*

See Appendix D for full details of the process and methods used to identify and select the clinical evidence relevant to the technology being appraised.

B.2.2 *List of relevant clinical effectiveness evidence*

Two Amgen-sponsored Phase II trials and a retrospective historical cohort were identified in the SLR that evaluated the treatment of adult MRD+ BCP-ALL patients in haematological CR:

- BLAST, a confirmatory, single-arm, open-label, international, multicentre study of 116 patients across 10 European countries[56]
- MT103-202, a pilot, single-arm, open-label, multicentre study of 21 patients in Germany[57]
- Historical cohort, a retrospective study of 182 patients designed to provide a well-matched cohort to the BLAST population[58]

Summaries of the blinatumomab clinical trials are provided in Table 5, Table 6, and Table 7 below.

Table 5: Clinical effectiveness evidence: BLAST

Study	BLAST (2014)[5, 56]		
Study design	Phase II, single-arm, open-label, international, multicentre		
Population	Adult MRD+ BCP-ALL patients in haematological CR after front-line therapy		
Intervention(s)	Blinatumomab 15 µg/m^2/day continuous infusion		
Comparator(s)	None		
Indicate if trial supports application for marketing authorisation	Yes	**Indicate if trial used in the economic model**	Yes
Rationale for use/non-use in the model	BLAST was used in the economic model as it is the primary study presented in this submission, and includes the largest population of MRD+ BCP-ALL patients receiving blinatumomab		
Reported outcomes specified in the decision problem	• Proportion of patients who achieved a complete MRD response within 1 cycle of blinatumomab • **Haematological RFS rate at 18 months* following initiation of blinatumomab** • **OS†** • Mortality rate within 100 days after allogeneic HSCT • TTHR • Duration of complete MRD response • Effect on MRD level • Overall incidence and severity of adverse effects • **Patient's quality of life during and after therapy (change from baseline in EORTC-QLQ-C30 and EQ-5D)**		
All other reported outcomes	NA		

Footnotes: *A patient-level analysis was used to assess RFS over the length of the trial follow-up in the economic analysis, and therefore also used additional timepoints. †A patient-level analysis was also performed to assess OS in the economic analysis.
Abbreviations: BCP: B-Cell Positive; ALL: Acute Lymphoblastic Leukaemia; CR: Complete Response; MRD: Minimal Residual Disease; RFS: Relapse-Free Survival; OS: Overall Survival; HSCT: Haematopoietic Stem Cell Transplantation; TTHR: Time to Haematological Relapse; EORTC-QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-C30; EQ-5D: EuroQol-5D.
Source: BLAST Key Secondary Analysis CSR[56]

Table 6: Clinical effectiveness evidence: Pilot

Study	MT103-202[57]		
Study design	Phase II, single-arm, open-label, multicentre		
Population	Adult MRD+ BCP-ALL patients in haematological CR after front-line therapy		
Intervention(s)	Blinatumomab 15 µg/m^2/day continuous infusion		
Comparator(s)	None		
Indicate if trial supports application for marketing authorisation	Yes	**Indicate if trial used in the economic model**	No
Rationale for use/non-use in the model	MT103-202 is not included in the economic model, as the confirmatory BLAST trial assessed blinatumomab in the same indication and a clinically similar population. MT103-202 is, however, included briefly in Section B.2.6.2 as a source of additional clinical data		
Reported outcomes specified in the decision problem	• Proportion of patients who achieved a complete MRD response within 4 cycles of blinatumomab • MRD response after any cycle • TTHR • MRD progression • MRD relapse after any cycle		
All other reported outcomes	NA		

Abbreviations: BCP: B-Cell Positive; ALL: Acute Lymphoblastic Leukaemia; CR: Complete Response; MRD: Minimal Residual Disease; RFS: Relapse-Free Survival; OS: Overall Survival; HSCT: Haematopoietic Stem Cell Transplantation; TTHR: Time to Haematological Relapse; EORTC-QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-C30; EQ-5D: EuroQol-5D.
Source: MT103-202 CSR[57]

Table 7. Clinical effectiveness evidence: Historical comparator

Study	20120148[58]		
Study design	Retrospective, single-arm, international, multicentre		
Population	Adult Ph-negative MRD+ BCP-ALL patients in haematological CR		
Intervention(s)	Standard of care chemotherapy regimens, according to national treatment or study group protocols		
Comparator(s)	None		
Indicate if trial supports application for marketing authorisation	No	**Indicate if trial used in the economic model**	Yes
Rationale for use/non-use in the model	Study 20120148 was designed to provide a well-matched patient population treated with SoC, and was compared to BLAST using propensity score matching to inform the economic model		
Reported outcomes specified in the decision problem	• Haematological RFS rate • OS		
All other reported outcomes	NA		

Abbreviations: BCP: B-Cell Positive; ALL: Acute Lymphoblastic Leukaemia; CR: Complete Response; MRD: Minimal Residual Disease; RFS: Relapse-Free Survival; OS: Overall Survival.
Source: Study 20120148 CSR[58]

MT103-202 was not used to populate the economic model but is included in sections 2.2 to 2.6. The results of this study support the findings presented for BLAST, which was a confirmatory study that followed the pilot MT103-202 study. This study was not included in the economic model because BLAST provides a much larger sample of MRD+ BCP-ALL patients treated with blinatumomab.

The historical comparator study was designed to provide a well-matched population of MRD+ patients treated with SoC, and was used to inform the comparative efficacy of blinatumomab using propensity score matching, as described in Section B.2.9.

B.2.3 *Summary of methodology of the relevant clinical effectiveness evidence*

B.2.3.1 Summary of trial methodology

Details of the BLAST and pilot (MT103-202) studies are presented in this section; the historical comparator study (20120148) is presented in detail in Section B.2.9.

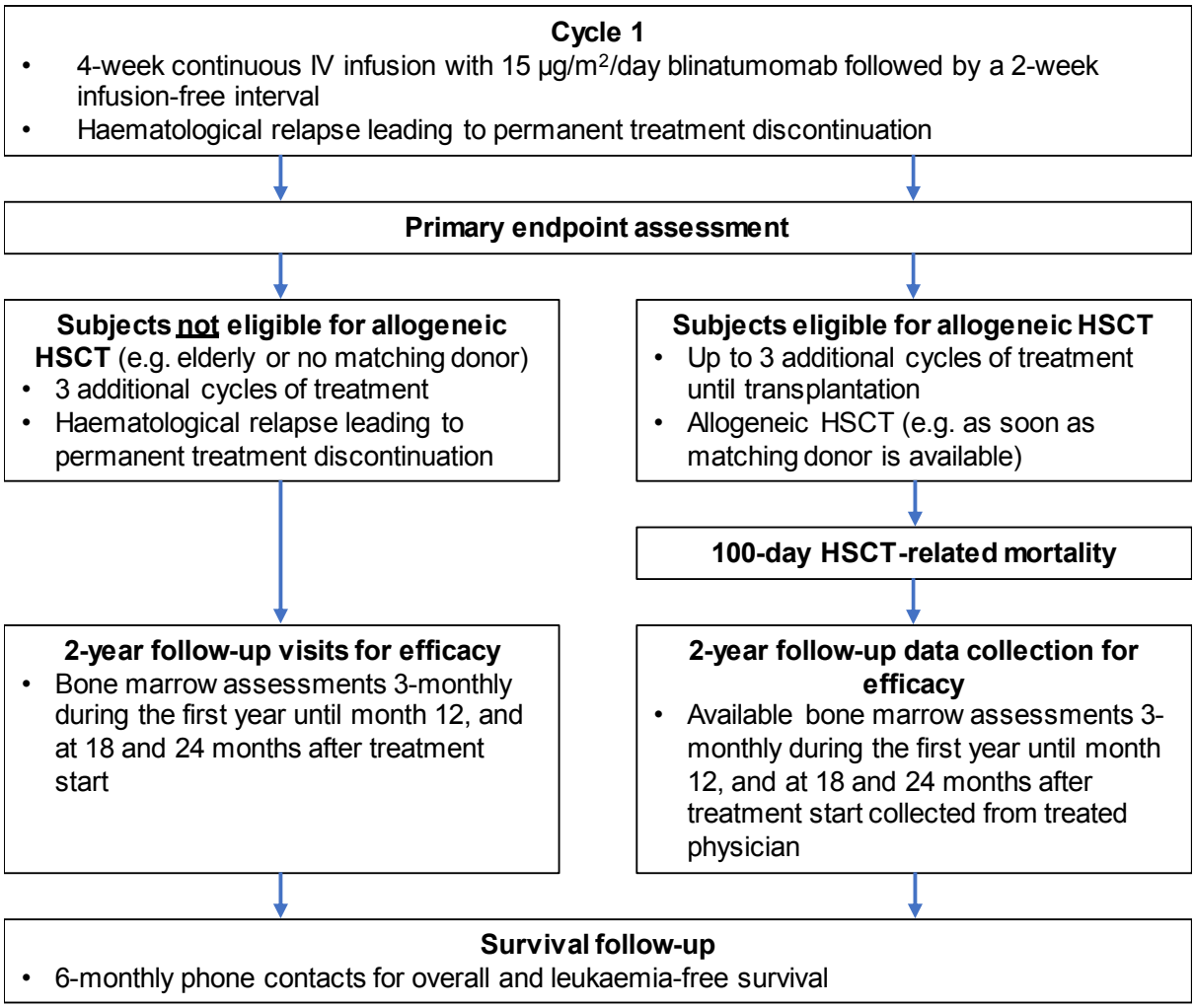
BLAST

BLAST was a Phase II, single-arm, open-label, multicentre trial to assess the efficacy, safety and tolerability of blinatumomab in adult patients with MRD+ BCP-ALL. An overview of the BLAST study design is presented in Figure 8. All patients in the study were intended to receive at least 1 and up to a maximum of 4 cycles of blinatumomab. A cycle was defined as a continuous intravenous (cIV) infusion at a constant dose of 15 µg m^{-2} day^{-1} over 4 weeks, followed by an infusion-free period of 2 weeks. The minimal criterion for inclusion in the Full Analysis Set (FAS),

which is consistent with the intention-to-treat principle in single-arm open-label studies, was 1 dose of blinatumomab.[56]

Upon completion of 1 cycle of treatment, all patients were assessed for the primary endpoint of MRD response rate. Those patients who were not candidates for allogeneic HSCT could continue treatment for up to 4 cycles; these patients were followed for efficacy, including bone marrow assessments, every 3 months for 2 years, then for survival follow-up every 6 months until 5 years after treatment start. Patients who were candidates for allogeneic HSCT could proceed to allogeneic HSCT immediately or after additional cycles of blinatumomab, for up to a maximum total of 4 cycles. For these patients, 100-day post-transplant mortality, 2-year efficacy and survival follow-up were assessed.[56]

Figure 8. Overview of study design for BLAST

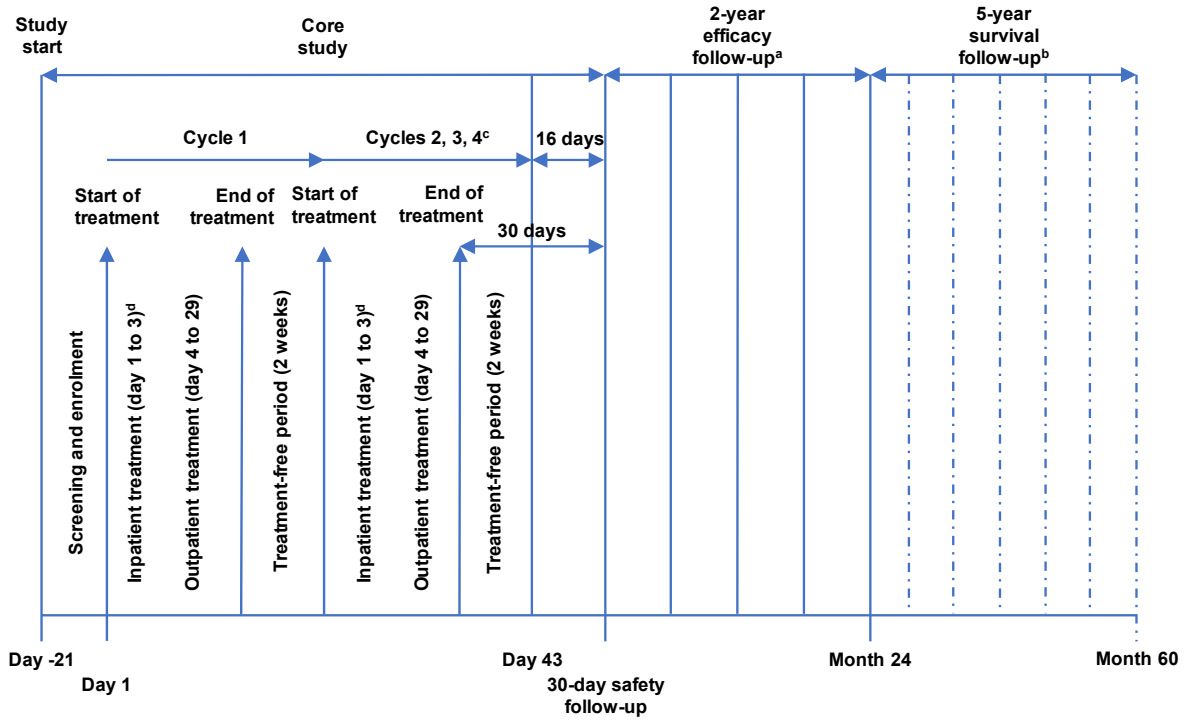


Abbreviations: IV: Intravenous; HSCT: Haematopoietic Stem Cell Transplantation.
Source: BLAST Key Secondary Analyses CSR Figure 8-1[56]

A schematic of the study design of BLAST in presented in Figure 9. Patients were treated for up to 4 cycles, unless haematological relapse occurred. A safety follow-up was performed 30 days after the end of the last infusion. Efficacy follow-ups occurred until 24 months after treatment start. After completion of the 2-year follow-up for haematological RFS, patients or their treating physicians were contacted by phone at least every 6 months for overall and leukaemia-free

survival follow-up until death or at least 5 years after treatment start, whichever occurred earlier.[56]

Figure 9. Schematic of study design for BLAST



Footnotes: [a]Efficacy follow-up visits at months 9, 12, 18, and 24 (± 2 weeks); [b]Survival follow-up visits by phone at months 30, 36, 42, 48, 54, and 60 (± 4 weeks); [c]Administration of up to 4 cycles of blinatumomab treatment at 15 µg/m^2/day, discontinuation of treatment due to haematological relapse; [d]Patients were hospitalised for at least 3 days after the start of treatment during cycle 1 and for at least 2 days after the start of treatment during subsequent cycles of treatment.
Source: BLAST Key Secondary Analyses CSR Figure 8-2[56]

Eligibility criteria

The inclusion and exclusion criteria for patients entering the BLAST study are listed in Table 8 below.

Table 8. Inclusion and exclusion criteria for patients in BLAST

	BLAST
Inclusion criteria	Patients were eligible for inclusion in the study only if all the following criteria applied: • Patients with B-precursor ALL in complete haematological remission defined as less than 5% blasts in bone marrow after at least three intense chemotherapy blocks (e.g., GMALL induction I–II/consolidation I, induction/intensification/consolidation or three blocks of Hyper CVAD) • Presence of MRD at a level of ≥10^{-3} (molecular failure or molecular relapse) in an assay with a sensitivity and a lower level of quantification of 10^{-4} documented after an interval of at least 2 weeks from last systemic chemotherapy • For evaluation of MRD, patients must have had at least one molecular marker based on individual rearrangements of immunoglobulin or TCR-genes or a flow cytometric marker profile evaluated by a national

	or local reference lab approved by the sponsor • Bone marrow specimen from primary diagnosis (enough DNA [30 pg] or a respective amount of cell material) for clone-specific MRD assessment must have been received by central MRD lab and lab must have confirm that the sample is available • Bone marrow function as defined below: o ANC (Neutrophils) ≥1,000/µL o Platelets ≥50,000/µL (transfusion permitted) o HB level ≥9g/dl (transfusion permitted) • Renal and hepatic function as defined below: o AST (GOT), ALT (GPT), and AP <2 x ULN o Total bilirubin <1.5 x ULN o Creatinine clearance ≥50 mL/min (calculated e.g. per Cockroft & Gault) • Negative HIV test, negative hepatitis B (HbsAg) and hepatitis C virus (anti-HCV) test • Negative pregnancy test in women of childbearing potential • ECOG Performance Status 0 or 1 • Age ≥18 years • Ability to understand and willingness to sign a written informed consent • Signed and dated written informed consent
Exclusion criteria	Patients were excluded from participation in the study if any of the following criteria applied: • Presence of circulating blasts or current extra-medullary involvement by ALL • History of relevant CNS pathology or current relevant CNS pathology (e.g. seizure, paresis, aphasia, cerebrovascular ischemia/haemorrhage, severe brain injuries, dementia, Parkinson's disease, cerebellar disease, organic brain syndrome, psychosis, coordination or movement disorder) • Current infiltration of cerebrospinal fluid by ALL • History of or active relevant autoimmune disease • Prior allogeneic HSCT • Eligibility for treatment with TKIs (i.e., Philadelphia chromosome-positive (Ph+) patients with no documented treatment failure of or intolerance/contraindication to at least 2 TKIs) • Systemic cancer chemotherapy within 2 weeks prior to study treatment (except for intrathecal prophylaxis) • Radiotherapy within 4 weeks prior to study treatment • Autologous HSCT within six weeks prior to study treatment • Therapy with monoclonal antibodies (rituximab, alemtuzumab) within 4 weeks prior to study treatment • Treatment with any investigational product within four weeks prior to study treatment • Previous treatment with blinatumomab • Known hypersensitivity to immunoglobulins or to any other component of the study drug formulation • History of malignancy other than ALL within five years prior to treatment start with blinatumomab, except for basal cell or squamous cell carcinoma of the skin, or carcinoma "in situ" of the cervix • Active infection, any other concurrent disease or medical condition

	that are deemed to interfere with the conduct of the study as judged by the investigator • Nursing women or women of childbearing potential not willing to use an effective form of contraception during participation in the study and at least 3 months thereafter or male patients not willing to ensure effective contraception during participation in the study and at least three months thereafter

Abbreviations: ALL: Acute Lymphoblastic Leukaemia; GMALL: German Multicentre ALL Working Group; CVAD: Cyclophosphamide, Vincristine, Doxorubicin (Adriamycin) and Dexamethasone; TCR: T-Cell Receptor; DNA: Deoxyribonucleic Acid; MRD: Minimal Residual Disease; ANC: Absolute Neutrophil Count; HB: Hemoglobin; AST: Aspartate Aminotransferase; GOT: Glutamic Oxaloacetic Transaminase; ALT: Alanine Aminotransferase; GPT: Glutamic-Pyruvic Transaminase; ULN: Upper Limit of Normal; HIV: Human Immunodeficiency Virus; HCV: Hepatitis C Virus Test; ECOG: Eastern Cooperative Oncology Group; CNS: Central Nervous System; HSCT: Haematopoietic Stem Cell Transplantation; TKI: Tyrosine Kinase Inhibitor.
Source: BLAST Study Protocol[59]

Settings and locations where the data were collected

BLAST was conducted in a secondary care (hospital) setting at 46 centres in Austria, Belgium, Czech Republic, France, Germany, Italy, Netherlands, Romania, Russia, Spain, and the United Kingdom. Seven patients (6.0%) were enrolled in the United Kingdom.[56]

Trial drugs and concomitant medications

A detailed overview of BLAST study drugs and required, permitted and disallowed concomitant medications is provided in Table 9 below.

Patients could have received up to 4 consecutive cycles of blinatumomab. A cycle consisted of a cIV infusion at a dose of 15 µg/m^2/day at a constant flow rate over 28 days followed by an infusion-free interval of 14 days, which could have been prolonged for up to 7 days, if necessary. Patients in haematological remission generally could receive up to 4 cycles of treatment, independently from achieving complete MRD response.[56] While a dosing schedule based on body surface area was used to evaluate blinatumomab in BLAST, a fixed dose regimen of 15 µg/m^2 day was found to result in similar drug exposure in blinatumomab trials for other indications, and the SmPC therefore recommends this fixed dose regimen for adults at least 45 kg in weight.[8]

In the case of neurologic (central nervous system) grade 3 adverse events, it was permitted to temporarily stop the treatment without discontinuing the study. If the event decreased to at least grade 1 within 1 week, treatment could be restarted again at a reduced blinatumomab dose of 5 µg/m^2/day within 2 weeks (but not earlier than 72 hours) after the infusion was stopped. In the case of interruption due to a clinically-relevant grade 2 neurological event, treatment could be restarted at either the original dose or the reduced dose of 5 µg/m^2/day after the adverse event decreased to at least grade 1, at the investigator's discretion. However, after dose reductions due to neurologic adverse events, re-escalation back to 15 µg/m^2/day was not permitted.[56]

If patients were suitable for allogeneic HSCT after treatment with at least 1 cycle of blinatumomab, they may have undergone allogeneic HSCT instead of receiving further cycles with blinatumomab. In the event of haematological relapse within the treatment period, treatment with blinatumomab was terminated.[56]

Table 9. Overview of BLAST study drugs and concomitant medications

Permitted concomitant medications	**Prior to the start of cycle 1:** • CSF prophylaxis (intrathecal triple combination) consisting of: dexamethasone 4 mg or equivalent, methotrexate 15 mg, cytosine arabinoside 40 mg administered during the screening period (prior to the baseline bone marrow assessment for MRD, unless this sequence was not feasible based on clinical considerations by the investigator) or within 4 weeks prior to study drug treatment, when done within clinical routine and at least consisting of methotrexate 15 mg • A corticosteroid (prednisone 100 mg IV or equivalent) administered within 1 hour before treatment start on day 1 **Prior to the start of subsequent cycles:** • A corticosteroid (prednisone 100 mg IV or equivalent) administered within 1 hour before treatment start on day 1 **During the treatment period:** • In case of neurologic events dexamethasone was administered orally at a dose of at least 24 mg/day for up to 3 days. The dose was then step-wise reduced over the next 4 days. If the neurologic event was a seizure, appropriate prophylactic anticonvulsant treatment with a therapeutic dose (e.g. phenytoin or levetiracetam) administered during restart and during start of the following new treatment cycle **Following treatment cycles 2 and 4 immediately after bone marrow aspiration (Day 29):** • A CSF prophylaxis consisting of an intrathecal triple combination regimen at absolute doses of 4 mg dexamethasone or equivalent, 15 mg methotrexate, and 40 mg cytosine arabinoside, unless this sequence was not feasible based on clinical considerations by the investigator **After completion of study treatment for patients who did not undergo HSCT:** • CSF prophylaxis was recommended every 3 months until at least month 18 and per the physician's discretion thereafter **For patients with a high risk for CMV infection (i.e. prior CMV reactivation), one of the following measures was performed:** • Intensive (twice weekly) CMV-PCR follow-up with early therapeutic intervention if positive, or prophylactic CMV treatment
Disallowed concomitant medications	**The following medication and therapies were prohibited during the study until end of efficacy period:** • Any anti-tumour therapy other than the investigational product: o Cytotoxic and/or cytostatic drugs o Radiation therapy o Immunotherapy • Any other investigational agent • Chronic systemic high-dose corticosteroid therapy (i.e. > 20 mg prednisone daily) • Any other immunosuppressive therapies (except for protocol mandated interventional corticosteroids) • Non-steroidal anti-inflammatory drugs (except for metamizole and/or naproxen), as they may affect the vascular system or, in case of acetylsalicylic acid, the platelet system. As naproxen can also affect the platelet system – although less pronounced and/or frequent than acetylsalicylic acid – it is second choice. Paracetamol/acetaminophen was allowed • Tyrosine kinase inhibitors

Abbreviations: CSF: Cerebrospinal Fluid; MRD: Minimal Residual Disease; IV: Intravenous; HSCT: Haematopoietic Stem Cell Transplantation; CMV: Cytomegalovirus; PCR: Polymerase Chain Reaction.
Source: BLAST Key Secondary Analysis CSR[56]

Outcomes used in the economic model or specified in the scope, including primary outcome

The pre-specified primary, key secondary and other secondary outcomes presented in this submission are provided in Table 10 below, with the outcomes that are used in the cost effectiveness analyses in Section B.3 highlighted in bold.

Table 10. Pre-specified primary, key secondary and other secondary outcomes from BLAST

	Outcome	Additional information
Primary outcome	• Proportion of patients who achieve complete MRD response defined by absence of MRD after one cycle of treatment with blinatumomab	• Complete MRD response was defined as no PCR amplification of individual rearrangements of Ig- or TCR-genes (the minimum required sensitivity of 1 x 10^{-4}) detected after completion of the first cycle
Key secondary outcome	• **Haematological relapse-free survival rate at 18 months* following initiation of blinatumomab**	• Haematological relapse was defined as unequivocal detection of > 5% leukaemia cells in bone marrow, presence of circulating leukaemia blasts, or extramedullary leukaemia (whichever occurs first)
Other secondary outcomes	• **OS†** • Mortality rate within 100 days after allogeneic HSCT • TTHR • Duration of complete MRD response • Effect on MRD level • Overall incidence and severity of adverse effects • **Patient's quality of life during and after therapy (change from baseline in EORTC-QLQ-C30 and EQ-5D)**	• MRD relapse was defined as the reappearance of individual rearrangements of Ig- or TCR-genes ≥LLOQ (usually 10^{-4}) for at least 1 individual marker measured by an assay with a sensitivity of minimum 10^{-4} in patients who had achieved complete MRD response

Footnotes: Outcomes highlighted in bold were included in the economic model. *A patient-level analysis was used to assess RFS over the length of the trial follow-up in the economic analysis, and therefore also used additional timepoints. †A patient-level analysis was also performed to assess OS in the economic analysis.
Abbreviations: MRD: minimal residual disease; PCR: Polymerase chain reaction; TCR: T cell receptor; OS: Overall survival; HSCT: haematopoietic stem cell transplant; TTHR: time to haematological relapse; EORTC-QLQ-C30: EORTC Quality of Life core questionnaire; EQ-5D: EuroQol five dimensions questionnaire; LLOQ: lower limit of quantification; RFS: relapse-free survival.
Source: BLAST Study Protocol[59]

MT103-202 (pilot study)

MT103-202 was designed to evaluate the efficacy and safety of blinatumomab in adult MRD+ BCP-ALL patients in haematological CR after front-line therapy. The study was conducted in collaboration with the German Multicenter Study group for Adult Acute Lymphoblastic Leukemia (GMALL) in Germany and is believed to be the first study conducted with an immunotherapy in this patient population.[57]

Adults (≥18 years of age) who were MRD+ at a level of at least $1x10^{-4}$ at any point after the first consolidation chemotherapy block (consolidation I) of front-line therapy were eligible for enrolment. Exclusion criteria included current extramedullary involvement, a history of (or current) clinically relevant central nervous system pathology, prior autologous HSCT (within 6 weeks) or allogeneic HSCT (at any time), or chemotherapy or radiotherapy (within 4 weeks). All eligible patients with a suitable donor were offered HSCT after the first blinatumomab cycle.[57]

The primary and secondary outcomes in MT103-202 are listed in Table 11 below.

Table 11. Pre-specified primary, key secondary and other secondary outcomes from MT103-202

	Outcome	Additional information
Primary outcome	• MRD response rate: o Incidence of MRD negativity/response within 4 cycles of treatment with blinatumomab	• If Ph+ or t(4;11), response achieved when Ph or t(4;11) was below detection limit and individual rearrangements of immunoglobulin or TCR genes are below 10^{-4}. If Ph and t(4;11) negative, response achieved when individual rearrangements of immunoglobulin or TCR genes are below 10^{-4}
Secondary outcomes	• MRD response after any cycle • TTHR • MRD progression • MRD relapse after any cycle	• TTHR: defined as > 5% leukaemia cells in bone marrow • Progression: the increase in the MRD level by 1 log as compared to the baseline level, which was equal to a 10-fold increase in the number of MRD cells • Relapse: reappearance of bcr/abl, and/or t(4;11) translocation at any detection level, and/or by individual rearrangements of immunoglobulin or TCR-genes ≥ 10^{-4} for ≥ 1 individual marker measured by an assay with a sensitivity of minimum 10^{-4}

Abbreviations: MRD: minimal residual disease; TCR: T cell receptor; RFS: relapse-free survival.
Source: MT103-202 CSR[57]

B.2.3.2 Comparative summary of trial methodology

A summary of the BLAST methodology, as well as the MT103-202 methodology, is included in Table 12.

Table 12. Comparative summary of trial methodology

Trial number (acronym)	MT103-203 (BLAST)	MT103-202 (Pilot)
Location	• 46 centres in Austria, Belgium, Czech Republic, France, Germany, Italy, Netherlands, Romania, Russia, Spain, and the United Kingdom	• 6 centres in Germany
Trial design	• Phase II, single-arm, open-label, international, multicentre	• Phase II, single-arm, open-label, multicentre
Eligibility criteria for participants	• Patients with BCP-ALL in haematological CR after at least 3 intense front-line chemotherapy blocks and presence of minimal residual disease at a level of ≥ 10^{-3}	• Patients with BCP-ALL in haematological CR after at least 3 intense front-line chemotherapy blocks and presence of minimal residual disease at a level of ≥ 10^{-3}
Settings and locations where the data were collected	• Secondary care (hospital) setting	• Secondary care (hospital) setting
Trial drugs	• Blinatumomab (n=116), cIV infusion at 15 µg/m^2/day at a constant flow rate over 28 days, followed by an infusion-free interval of 14 days, for up to 4 cycles	• Blinatumomab (n=21, cIV infusion at 15 µg/m^2/day at a constant flow rate over 28 days, followed by an infusion-free interval of 14 days, for up to 7 cycles
Permitted and disallowed concomitant medication	**Permitted medications** • Prior to the start of cycle 1: o CSF prophylaxis o A corticosteroid • Prior to the start of subsequent cycles: o A corticosteroid • During the treatment period: o Dexamethasone in the case of neurologic events • Following treatment cycles 2 and 4 immediately after bone marrow aspiration: o CSF prophylaxis • After completion of study treatment for patients who did not undergo HSCT: o CSF prophylaxis	**Permitted medications** • Premedication for each treatment cycle included a corticosteroid to suppress cytokine release (100 mg methylprednisolone IV at 1 hour prior to start of blinatumomab infusion or prior to restart if infusion interruption > 12 hours) and thrombosis prophylaxis by low molecular weight heparin (subcutaneous) during the first 7 days of each treatment cycle • CNS prophylaxis was administered with the following intrathecal triple combination regimen at absolute doses: dexamethasone 4 mg, methotrexate 15 mg, cytosine-arabinoside 40 mg. If the patient had MRD response after cycle 1 of treatment, the triple combination regiment was administered immediately after the first bone marrow aspiration study on day 28 of cycle 2 • In non-responders, after cycle 1 demonstrated detectable MRD, the triple combination regimen was

	• Patients at high risk for CMV infection: o Intensive CMV-PCR follow-up or prophylactic CMV treatment **Disallowed medications** • Any anti-tumour therapy • Any other investigational agent • Chronic systemic high-dose corticosteroid therapy • Any other immunosuppressive therapies • Non-steroidal anti-inflammatory drugs • Paracetamol/acetaminophen was allowed • Tyrosine kinase inhibitors	administered after cycle 3 of treatment immediately after bone marrow aspiration on cycle day 28 of cycle 3. CNS prophylaxis continued every 3 months • Small molecule tyrosine kinase inhibitors registered for the treatment of ALL disease were permitted as concomitant treatment of patients with bcr/abl positive MRD if the patients developed MRD relapse on tyrosine kinase inhibitors or whose MRD persisted on tyrosine kinase inhibitors for more than 8 weeks • For symptomatic treatment of fever, metamizole was administered **Disallowed medications** • Any anti-tumour therapy other than blinatumomab as indicated in the protocol • Any other investigational agent • Chronic systemic high-dose corticosteroid therapy • Other immunosuppressive therapies • Stem-cell transplantation • Any use of NSAIDs (except for paracetamol)
Primary outcomes (including scoring methods and timings of assessments)	• MRD response rate: the proportion of subjects who achieved a complete MRD response rate defined by the absence of MRD within 1 cycle of treatment with blinatumomab	• MRD response rate: the incidence of MRD negativity/response within 4 cycles of treatment with blinatumomab
Other outcomes used in the economic model/specified in the scope	• Haematological relapse-free survival rate at 18 months following initiation of blinatumomab • OS • Mortality rate within 100 days after allogeneic HSCT • TTHR • Duration of complete MRD response • Effect on MRD level • Overall incidence and severity of adverse effects • Patient's quality of life during and after therapy (change from baseline in EORTC-QLQ-C30 and EQ-5D)	• MRD response after any cycle • TTHR • MRD progression • MRD relapse after any cycle

Pre-planned subgroups	Subgroup analyses were performed to determine the effect of the following baseline covariates on MRD response, RFS, OS, TTHR and HSCT: • Age (15 to 34 years, 35 to 54 years, 55 to 64 years, and ≥ 65 years) • Gender • By Ph+ disease • t(4;11) translocation and/or MLL-AF4+ ALL • First, second, and further haematological remission • MRD level at baseline (< 1 x 10^{-2} versus ≥ 1 x 10^{-2}) • White blood cell (WBC) count at first diagnosis ≤ 30,000/mL and > 30,000/mL • Prior treatment regimen for ALL (type of therapy and, if applicable the drug-name) • Chemoresistance after the first week of chemotherapy • Need of a second induction course (salvage) for complete haematological remission • Previous anti-tumour radiotherapies • Haploid or near-triploid ALL • Clinical trial material from manufacturing processes 4 or 5 (CTM4 versus CTM5)	Subgroup analyses were performed to determine the effect of the following baseline covariates on the primary and secondary outcomes: • Primary outcome: o Baseline MRD assessment o Rearrangements (immunoglobulin or TCR genes) and translocations (bcr/abl and/or t[4;11] genes) • TTHR: o HSCT status • Other secondary endpoints: o Baseline MRD assessment o Rearrangements (immunoglobulin or TCR genes) and translocations (bcr/abl and/or t[4;11] genes)

Abbreviations: BCP-ALL: B-cell precursor acute lymphoblastic leukaemia; CTM4/5: clinical trial material from manufacturing process 4/5; ALL: acute lymphoblastic leukaemia; WBC: white blood cell; MRD: minimal residual disease; HSCT: haematopoietic stem cell transplant; TTHR: time to haematological relapse; OS: Overall survival; RFS: relapse-free survival; CMV-PCR: cytomegalovirus polymerase chain reaction.
Source: BLAST Key Secondary Analysis CSR[56] and BLAST Protocol[59]

B.2.3.3 Baseline characteristics

Baseline characteristics of participants in BLAST are described in Table 13. Overall, 116 patients with MRD+ BCP-ALL were enrolled into BLAST and comprised the FAS. The median age was 45 years; 13% of patients were 65 years of age or older. Nearly two-thirds (65%) of patients were in first haematological CR.[56]

Table 13. Baseline characteristics of participants in BLAST

Baseline characteristic	MT103-203 (BLAST) (n=116) n (%)
Male sex, n (%)	68 (59)
Median age (range), years	45.0 (18–76)
Age, n (%)	
≥18 to <35 years	36 (31.0)
≥35 to <55 years	41 (35.3)
≥55 to <65 years	24 (20.7)
≥65 years	15 (12.9)
Median time from prior treatment (range), months	1.3 (0–45)
Relapse history, n (%)	
First CR	75 (65)
Second CR	39 (34)
Third CR	2 (2)
Baseline MRD levels, n (%)	
$\geq 10^{-1} < 1$	9 (7.8)
$\geq 10^{-2} < 10^{-1}$	45 (38.8)
$\geq 10^{-3} < 10^{-2}$	52 (44.8)
$< 10^{-3}$	3 (2.6)
Below LLQ	5 (4.3)
Unknown	2 (1.7)
Philadelphia chromosome disease status	
Positive	5 (4.3)
Negative	111 (95.7)

Abbreviations: CR: Complete response; LLQ: lower limit of qualification.
Source: Adapted from Pharmaceutical Benefits Advisory Committee (2008) Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Canberra: Pharmaceutical Benefits Advisory Committee. BLAST Key Secondary Analysis CSR[56]

Baseline characteristics of participants in MT103-202 are presented in Table 14. Most patients (60%; 12/20) were female, Caucasian (100%; 20/20), with translocations of rearrangements of immunoglobulin/TCR genes (65%; 13/20). Overall, 45% (9/20) of patients were > 60 years of age.[56]

Table 14. Baseline characteristics of participants in MT103-202

Baseline characteristic	MT103-202 (Pilot) (n=20) n (%)
Age (years) **20-30**	3 (15.0)
31-40	5 (25.0)
41-50	2 (10.0)
51-60	1 (5.0)
61-70	7 (35.0)
> 70	2 (10.0)
Sex **Male**	8 (40.0)
Female	12 (60.0)
Race **Caucasian**	20 (100.0)
Translocations (all)[a] **bcr/abl above detection limit (all)**	[redacted]
t(4;11) translocation above detection limit (all)	2 (10.0)
Rearrangements of immunoglobulin/TCR genes (only)[b]	[redacted]
Rearrangements of immunoglobulin/TCR genes and translocations[c]	[redacted]

Footnotes: [a]Patients may have rearrangements in addition to translocations; [b]Patients did not show any translocation; cPatients showed rearrangements and translocations.
Abbreviations: bcr/abl: breakpoint cluster region; TCR: T-cell receptor.
Source: Adapted from Pharmaceutical Benefits Advisory Committee (2008) Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Canberra: Pharmaceutical Benefits Advisory Committee. MT103-202 CSR[57]

B.2.4 *Statistical analysis and definition of study groups in the relevant clinical effectiveness evidence*

A summary of the analysis populations for efficacy and safety outcomes for both BLAST and MT103-202 is presented in Table 15, while a summary of statistical analyses for the primary efficacy analysis in the trials is presented in Table 16.

Details of the participant flow for the two blinatumomab trials are presented in Appendix D.

Table 15. Summary of analysis populations

	BLAST	MT103-202
Primary efficacy analysis	All patients who received any infusion of blinatumomab with	All patients in the FAS.

	an Ig or TCR PCR MRD assay with the minimum required sensitivity of 1 x 10^{-4}, measured at a central lab, established at baseline (the primary endpoint full analysis set [Prim EP FAS]). This definition is consistent with the intention-to-treat principle in single-arm open-label studies.	
Key secondary analysis	Patients in the FAS who were in haematological CR at treatment start, excluding Philadelphia-positive patients.	All patients in the FAS.
Other secondary analyses	All patients in the FAS.	All patients in the FAS.
Safety analysis	All patients in the FAS.	All patients in the FAS.

Abbreviations: TCR: T-Cell Receptor; PCR: Polymerase Chain Reaction; MRD: Minimal Residual Disease; EP: Endpoint; FAS: Full Analysis Set; CR: Complete Remission.
Source: BLAST Key Secondary Analysis CSR[56], MT103-202 CSR[57]

Table 16. Summary of statistical analyses for the primary efficacy analysis in BLAST and MT103-202

Trial number (acronym)	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
MT103-203 (BLAST)	• The primary efficacy endpoint of the study was the proportion of patients who achieved a complete MRD response rate defined by the absence of MRD within 1 cycle of treatment with blinatumomab. • The following hypotheses were tested in this study: o H_0: $\pi \leq p_0 = 44\%$ versus H_1: $\pi \geq p_1 = 61\%$ • The following assumptions were made for the statistical hypothesis of the study: p_0, the MRD response probability, which, if true, means that blinatumomab was not worth studying further, was estimated to be not higher than 44%. • The future use of blinatumomab would be of considerable interest if the true MRD response probability (π) was 61% or higher (p_1).	• Analysis was performed by calculating the response rate and the 2-sided exact 95% CI and by covariates. • The complete MRD response rate is calculated as: o Number of patients with MRD- CR after 1 cycle of treatment divided by all patients.	• 100 patients were required, with a 90% power of demonstrating that the 97.5% 1-sided exact CI that the primary endpoint excluded 44% (p_0) if the true unknown response rate was 61% (p_1). • If the study observed at least 55 out of 100 patients (55%) with a complete MRD response after one cycle of treatment, then the null hypothesis (H_0) was rejected. • For this study, the recruitment rate was higher (n=116). In the case that more than 100 evaluable patients were recruited, the following parameters were adjusted for the primary efficacy endpoint: o N = 110 patients: H_0 could be rejected with 60/110 (= 55%) of MRD− patients. o N = 120 patients: H_0 could be rejected with 64/120 (= 53%) of MRD− patients.	• Only non-missing data were analysed, missing clinical data were not replaced. • Patients withdrawn prior to the end of the first cycle of blinatumomab treatment or later were not replaced. • It is recognised that the definition of the FAS, which excludes – for the analysis of the primary efficacy endpoint – treated patients for whom no sufficient MRD assessment by PCR could be established due to technical reasons, stretches the concept of 'intention to treat'. • However, this definition avoids the necessity to impute certain missing assessments as either 'no events' (best case) or as 'events' (worst case). • It seems highly plausible that the probability of the exclusion of an assessment/ patient from the primary efficacy analysis due to technical reasons with the PCR assay is uncorrelated to any patient characteristics in this indication, therefore the underlying missing data

				pattern would be 'missing completely at random.[60]
MT103-202 (pilot)	• The following hypotheses were tested in this study: o H_0: $\pi \leq p_0 = 5\%$ versus H_1: $\pi \geq p_1 = 30\%$. • The following assumptions are made for the statistical hypothesis of the study: P_0, the MRD response probability, which, if true, means that the agent was not worth studying further, was estimated to be not higher than 5%. The future use of blinatumomab would be of considerable interest if the true MRD response probability (π) was 30% or higher (p_1).	• Exact 2-sided 95% Clopper-Pearson confidence intervals (CIs) for the MRD response rate in each cohort and overall were provided. • P-values from the 1-sided exact binomial test for H_0 were provided in addition.	• General considerations were based on Simon's 2-Stage MinMax design with the following specifications: $p_0 = 0.05$, $p_1 = 0.3$, alpha = 0.05, power = 0.8. • The sample size was calculated as detailed below: o 7 patients were planned to be accrued during stage 1. If "0" responses were observed during stage 1, then the study would have been stopped after stage 1. o Otherwise the sample size would increase to 14 and the criteria for success/failure would have been: failure if ≤ 2/14 responses were observed and so no further investigation of the drug was warranted, success if o ≥ 3/14 responses were observed. • When the DRC met to review the data from the first 4 patients enrolled in the study, the threshold for declaring success was already reached with 3/4 responses. The DRC recommended: o Dose increase after cycle 1 for the non-responders	• Treatment with blinatumomab was discontinued in the event of any of the following: o Haematological relapse o MRD relapse o Progressive disease of MRD o Investigator's decision that a change of therapy was in the patient's best interest, in particular when a stem cell donor became available o Withdrawal of patient's consent o Patient or investigator not compliant with the study protocol o Progression of a medical condition which in the opinion of the investigator precluded further participation of the patient in the study o Administration of non-permitted concomitant medication(s) o Occurrence of an adverse event which made discontinuation desirable or

			○ Keeping the Simon's 2-stage MinMax design unchanged and then enrolling 7 more patients to reach a total sample size of 21 in order to obtain more data in safety and efficacy. ○ The first 4 patients enrolled in the run-in dose finding cohort were considered as part of the stage 1 part of the Simon's 2-stage design. Thus, the original protocol design was amended accordingly on 27 October 2008.	○ necessary in the investigator's and/or the patient's opinion. • All reasons for treatment discontinuation were clearly and concisely documented in the electronic case report form (CRF). If a patient had not continued to present him/herself during the study, the investigator was to describe the reason and circumstances as completely and accurately as possible. • Patients who terminated the study before the end of the first treatment cycle were not assessable regarding efficacy and were to be replaced.

Abbreviations: MRD: minimal residual disease; CR: Complete response.
Source: BLAST Key Secondary Analysis CSR[56], MT103-202 CSR[57]

B.2.5 *Quality assessment of the relevant clinical effectiveness evidence*

In order to assess the risk of bias and generalisability of the relevant clinical effectiveness trials, quality assessment was conducted using the Cochrane risk of bias checklist. The quality of non-randomised studies was assessed using the ROBINS-I checklist. These quality assessment checklists are included in Appendix D.

Whilst BLAST and the pilot study were open label, single-arm trials, the studies were well conducted and a low risk of bias was detected.

B.2.6 *Clinical effectiveness results of the relevant trials*

Summary of Clinical Effectiveness Results

- Blinatumomab is the only therapy specifically indicated for adults with BCP-ALL in haematological CR that can be used to achieve MRD negativity.
- BLAST (n=116) demonstrates the ability of blinatumomab to induce a complete MRD response in 78% of patients with MRD+ BCP-ALL in haematological CR.
- Patients who achieved complete MRD response with blinatumomab had 17.9 months longer median haematological RFS than those who did not.
- Patients treated with blinatumomab in first haematological remission had more than double the median haematological RFS of those in later remissions, and patients experiencing a complete MRD response with blinatumomab had longer RFS and OS than non-responders.
- Patients who achieved complete MRD response within 1 cycle of treatment of blinatumomab had a median OS more than triple that of patients who did not.
- Furthermore, blinatumomab offers durable complete MRD responses, with a median duration of 17.3 months.
- The highly effective results of blinatumomab treatment were also reflected in the pilot Phase II trial (MT103-202; n=20):
  - 80% of evaluable patients achieved MRD response, with all MRD responses having been observed within the first treatment cycle.
  - Median duration of complete MRD response for patients was 13.0 months.
  - Median haematological RFS had not been reached after a median follow-up time of 1550 days (> 4 years).
  - After up to 2138 days (more than 5 years) follow-up 52.6% of patients treated with blinatumomab remained relapse-free.
- Taken together, these results suggest that blinatumomab can achieve MRD negativity in most patients, providing long-lasting benefits to OS and RFS.

B.2.6.1 BLAST

Overview of data presentation

The analyses presented in this submission were conducted after the last Ph-negative patient completed an 18-month follow-up period (as Ph-positive patients were excluded from the pre-specified secondary analyses, as described in Section B.2.4), with the data cut-off date of 5th August 2015. All BLAST pre-specified primary and secondary efficacy endpoints are presented in detail in the main submission as all are relevant to the decision problem and included in the final

scope for this appraisal, as presented in Table 17. The FAS (n=116), which is consistent with the ITT principle in single-arm, open label studies, was used throughout the presented analyses.

Table 17. Overview of BLAST clinical effectiveness results presented in the main submission

Pre-specified primary endpoint	• Complete MRD response within 1 cycle
Pre-specified secondary endpoints	Key pre-specified secondary endpoint • Haematological RFS rate Other pre-specified secondary endpoints • OS • Mortality rate within 100 days after allogeneic HSCT • TTHR • Duration of complete MRD response • Effect on MRD level • Change from baseline EORTC-QLQ-C30 • Change from baseline EQ-5D

Abbreviations: MRD: Minimal Residual Disease; RFS: Relapse-Free Survival; OS: Overall Survival; HSCT: Haematopoietic Stem Cell Transplant; TTHR: Time To Haematological Relapse; EORTC-QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-C30; EQ-5D: EuroQol-5D.
Source: BLAST Key Secondary Analysis CSR[56]

Primary outcome (complete MRD response within 1 cycle)

- **Blinatumomab achieves MRD negativity in 78% of MRD+ BCP-ALL patients in haematological CR within the first cycle.**

The primary outcome presented for BLAST is the proportion of patients who achieved a complete MRD response within 1 cycle of blinatumomab, which was achieved by 77.9% of patients (Table 18). The MRD response rate was significantly greater than the null hypothesis threshold for the study of 44%.[56]

Table 18. Proportion of patients achieving complete MRD response within 1 cycle

Outcome	Blinatumomab
N	113
Response rate, n (%)	88 (77.9)
95% CI, %	69.1, 85.1

Abbreviations: MRD: Minimal Residual Disease; CI: Confidence Interval.
Source: BLAST Key Secondary Analysis CSR Table 14-4.1.1[56]

After the first treatment cycle, 2 additional patients had a complete MRD response, resulting in an overall complete response of 79.6%. For these 90 patients, median time to MRD response was 29.0 days, as presented in Table 19.[56]

Table 19. Overall MRD response

Outcome	Blinatumomab
N	113
Response rate, n (%)	90 (79.6)

95% CI, %	71.0, 86.6
Time to MRD response, N	90
Median (range)	29.0 (5–71)

Abbreviations: MRD: Minimal Residual Disease; CI: Confidence Interval.
Source: BLAST Key Secondary Analysis CSR Table 14-4.1.1[56]

Secondary outcomes

In the following sections, the secondary analyses performed in BLAST are described in detail. Results are presented both with and without censoring at HSCT or post-blinatumomab chemotherapy. It is important to note that it is not appropriate to compare the RFS and OS of patients who received HSCT with those who did not, for the following reasons:

- Kaplan-Meier (KM) curves and rates for transplanted patients are positively biased as they must have lived long enough to receive HSCT, resulting in an artificial plateau on the KM curve. In contrast, non-transplanted patients are negatively biased as patients receiving HSCT are later excluded, creating an artificial crash on the KM curve.

- The death rate in ALL is particularly high during the first 3 months of treatment, but reaches a plateau shortly after this point; as patients typically receive HSCT after a wait of several months, even if the transplant had no effect an artificially higher KM curve would be observed.

- It is plausible that patients who undergo transplantation may have different characteristics compared to those who do not undergo transplantation. For instance, less fit patients may not be eligible for HSCT, and the patients with the worst prognosis at baseline may not survive long enough to find a donor and receive HSCT. As such, transplanted and non-transplanted patients are not compared in these analyses.

Nonetheless, the analyses with censoring at HSCT or post-blinatumomab chemotherapy do suggest that blinatumomab provides improved outcomes independent of HSCT status.[58]

Haematological RFS at 18 months

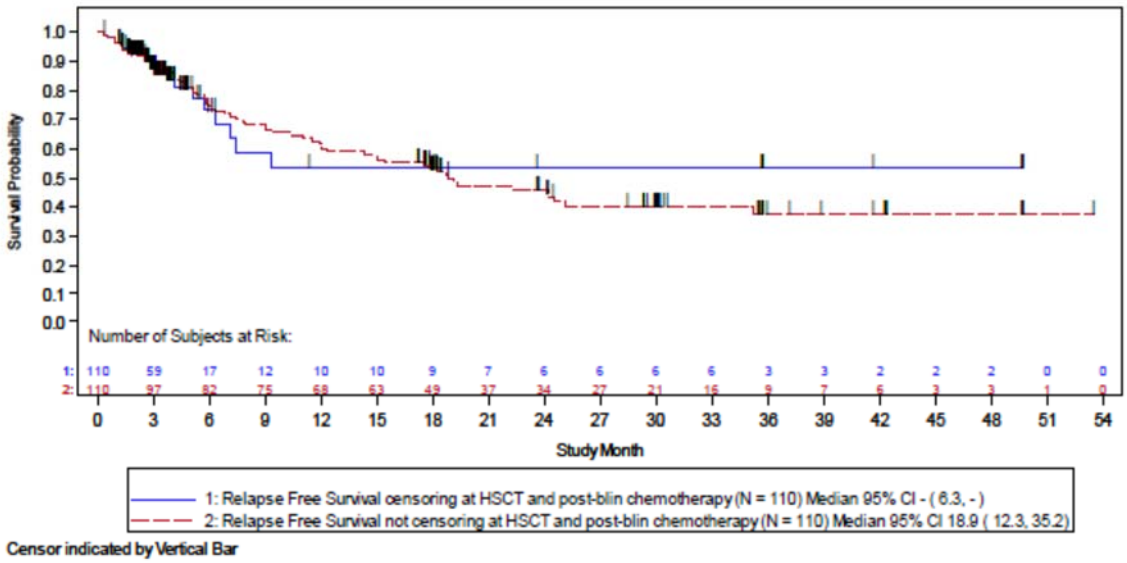
- **Haematological RFS at 18 months was a clinically-meaningful 54%, while median RFS was not estimable after more than 40 months.**
- **Patients who achieved complete MRD response within 1 cycle of treatment with blinatumomab had 17.9 months longer median haematological RFS than those who did not.**

The key secondary outcome was the haematological RFS rate at 18 months in all Ph− patients, censoring at HSCT or post-blinatumomab chemotherapy.[56]

As presented in Figure 10 and Table 20, the rate of haematological RFS at 18 months with censoring was 54% (95% CI: 33%, 70%); as the 33% lower boundary of the 95% CI exceeded the pre-specified threshold of 28%, the 18-month haematological RFS rate was clinically meaningful. These results demonstrate that blinatumomab provides improved outcomes independent of transplant status. The median RFS was not estimable at the time of data cut-off (more than 40 months). Without censoring, the 18-month haematological RFS rate was 53% (95% CI: 44%, 62%). It should be noted that while the K-M curve suggests improved RFS for

non-transplanted patients, this analysis is limited by the low number of patients included, with only 10 patients included from 12 months onwards.[56]

Figure 10. RFS with and without censoring at allogeneic HSCT and post-BLINCYTO chemotherapy



Abbreviations: RFS: relapse-free survival; HSCT: haematopoietic stem cell transplant.
Source: BLAST Key Secondary Analysis CSR Figure 14-4.2.1[56]

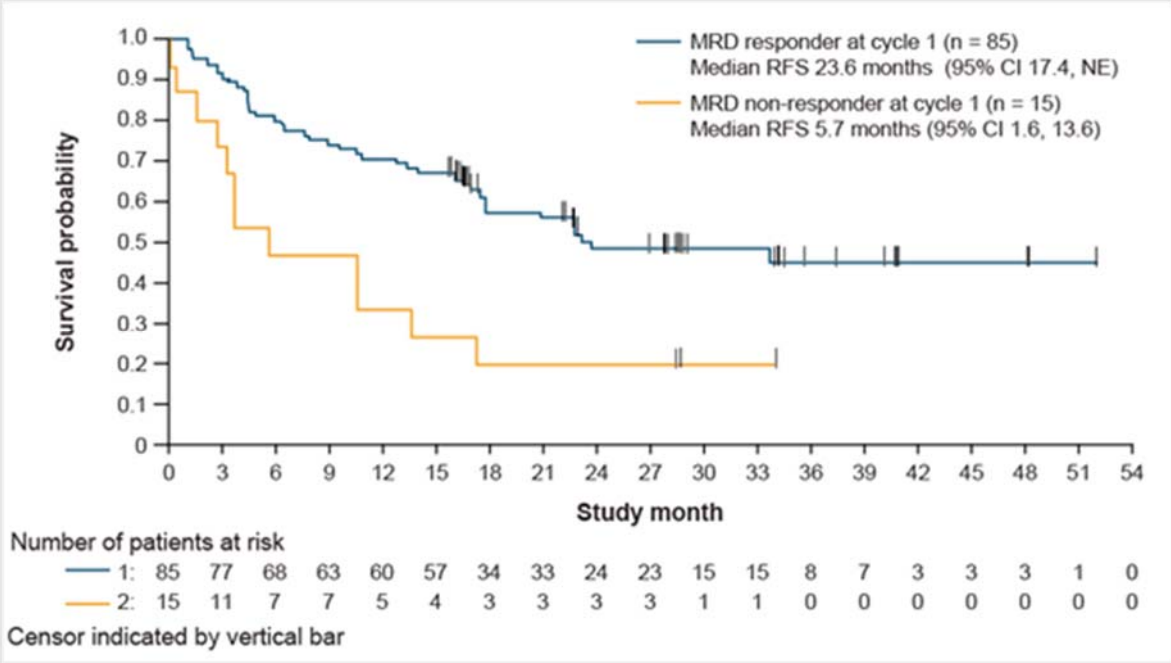
Table 20. RFS at 18 months with and without censoring at allogeneic HSCT and post-BLINCYTO chemotherapy

Outcome	Blinatumomab (N=110)	Blinatumomab censored (N=110)
Events, n (%)	62 (56.4)	21 (19.1)
Censors, n (%)	48 (43.6)	89 (80.9)
RFS (18 months)	0.53	0.54
95% CI	0.44, 0.62	0.33, 0.70

Abbreviations: RFS: relapse-free survival; HSCT: haematopoietic stem cell transplant; CI: Confidence Interval.
Source: BLAST Key Secondary Analysis CSR Table 14-4.2.1[56]

Furthermore, patients who achieved a complete MRD response to blinatumomab in cycle 1 achieved a statistically significantly (p=0.003) greater haematological RFS at 18 months than those who did not respond (58% versus 20%), as presented in Figure 11. The patients who achieved a complete MRD response in one cycle of blinatumomab experienced a median haematological RFS of approximately 18 months longer than those who did not (23.6 months versus 5.7 months).[56]

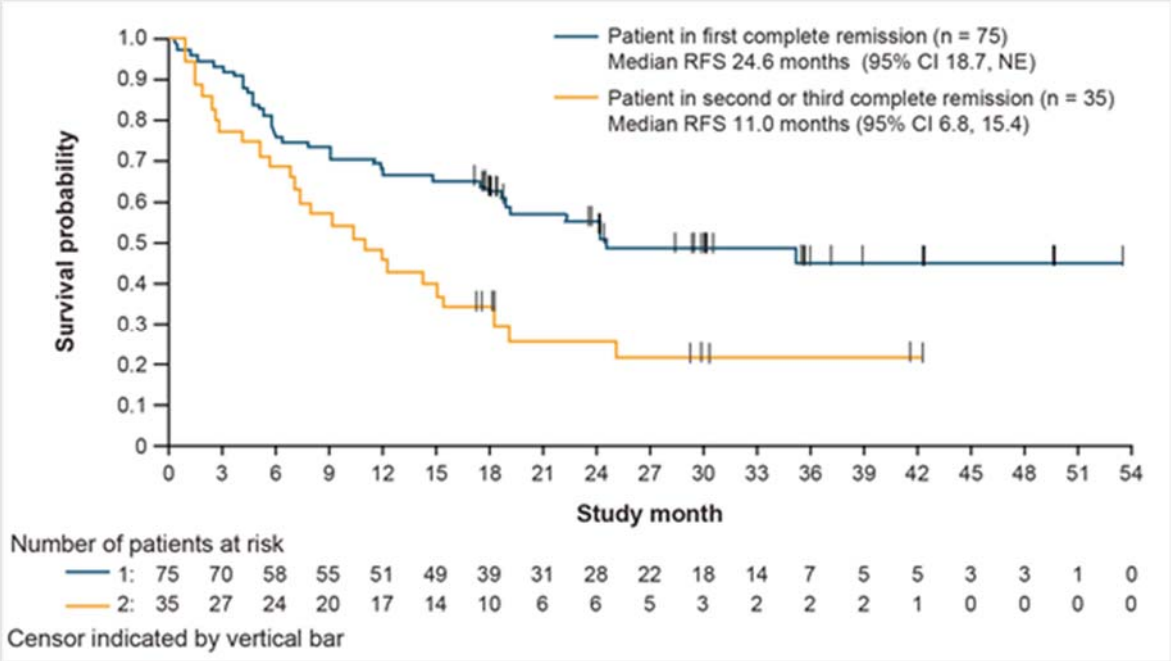
Figure 11. RFS in patients with or without a complete MRD response in cycle 1



Abbreviations: RFS: relapse-free survival; MRD: minimal residual disease.
Source: Gökbuget et al. (2015)[61]

Relapse history was also statistically significantly associated with haematological RFS (p=0.004), as shown in Figure 12. Median haematological RFS for patients in their first haematological CR was more than double that for patients in their second or third CR (24.6 versus 11.0 months), as was the 18-month haematological RFS (62% versus 34%).[56]

Figure 12. RFS in patients in first haematological remission or second/third remission



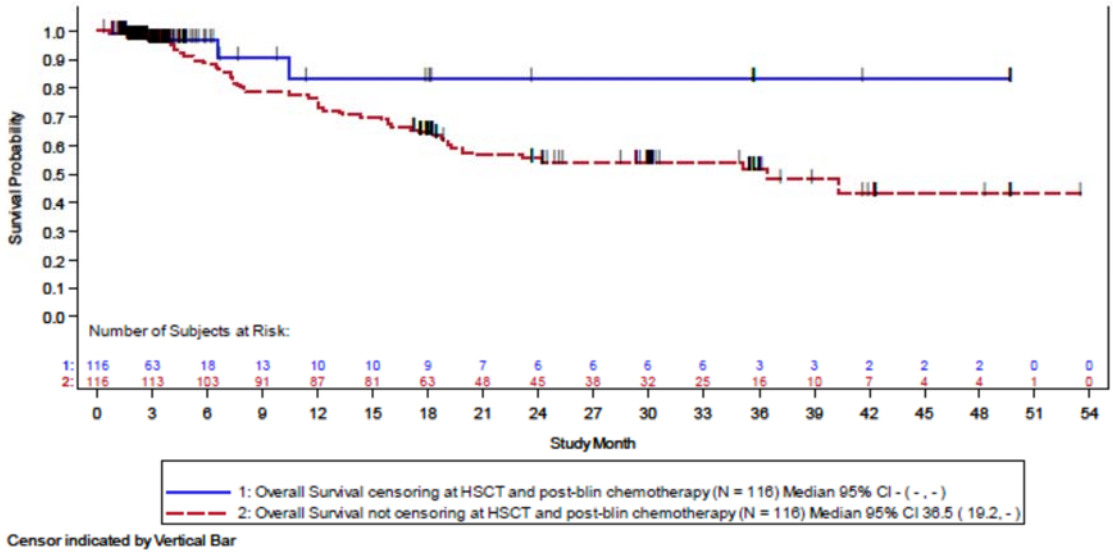
Abbreviations: RFS: relapse-free survival.
Source: Gökbuget et al. (2015)[61]

Overall survival

- **OS at 18-months was 65%, with a median OS of 36.5 months; with censoring at HSCT or post-blinatumomab chemotherapy, 18-month OS was 83% and median OS was not estimable after more than 40 months.**
- **Patients who achieved complete MRD response within 1 cycle of treatment of blinatumomab had a median OS more than triple that of patients who did not.**

Overall survival was measured for all patients from the time that they first received blinatumomab until death due to any cause, with patients who did not die censored at their last contact date. A total of 53 deaths (45.7%) were reported in the study as of the cut-off date. The 18-month OS in BLAST was 65% (95% CI: 55%, 73%) and median OS was 36.5 months (95% CI: 19.8%, not estimable), as presented in Figure 13 and Table 21. A sensitivity analysis for censoring at HSCT or post-blinatumomab chemotherapy is also included in Figure 13 and Table 21, with an 18-month OS of 83% (95% CI: 55%, 94%), suggesting that blinatumomab provides improved outcomes independent of transplant status. The median OS with censoring was not estimable. It should be noted that while the K-M curve suggests improved OS for non-transplanted patients in comparison to transplanted patients, this analysis is limited by the low number of patients included, with only 10 patients included from 12 months onwards.[56]

Figure 13. OS in BLAST with and without censoring at allogeneic HSCT and post-BLINCYTO chemotherapy



Abbreviations: OS: Overall survival; HSCT: haematopoietic stem cell transplant.
Source: BLAST Key Secondary Analysis CSR Figure 14-4.3.1[56]

Table 21. OS in BLAST at 18 months with and without censoring at allogeneic HSCT and post-BLINCYTO chemotherapy

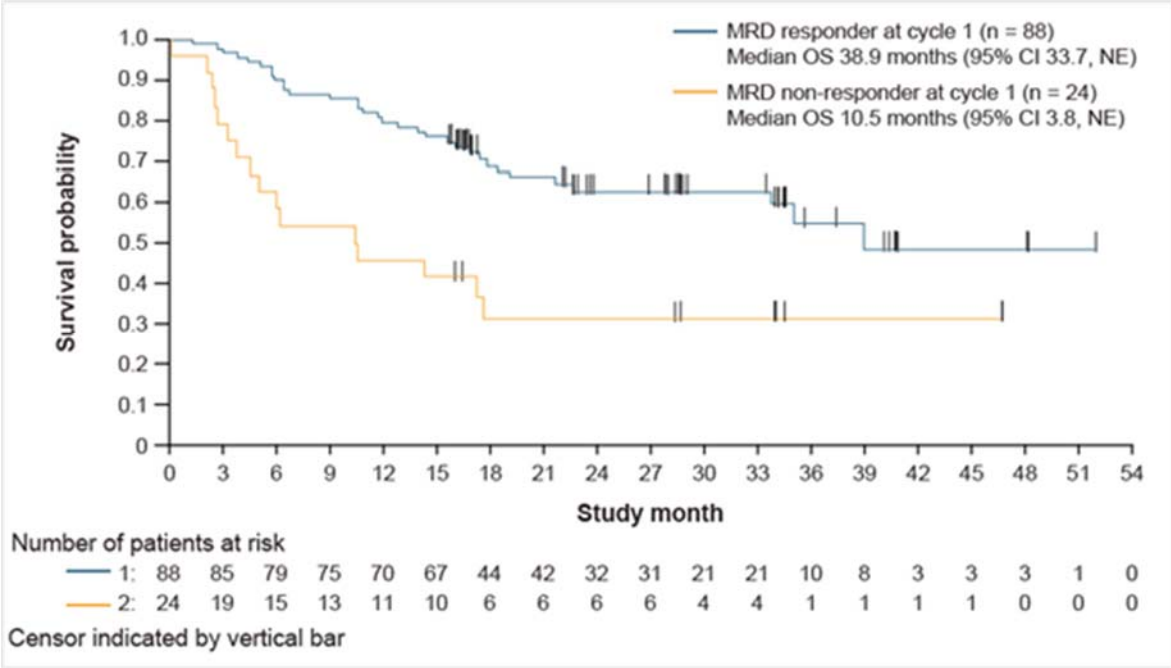
Outcome	Blinatumomab (N=116)	Blinatumomab censored (N=116)
Events, n (%)	53 (45.7)	5 (4.3)
Censors, n (%)	63 (54.3)	111 (95.7)
OS (18 months)	0.65	0.83

95% CI	0.55, 0.73	0.55, 0.94
Median (95% CI)	36.5 (19.2, n.e.)	n.e. (n.e., n.e.)

Abbreviations: OS: Overall survival; HSCT: haematopoietic stem cell transplant; CI: Confidence Interval; n.e.: not estimable.
Source: BLAST Key Secondary Analysis CSR Table 14-4.3.1[56]

As for the primary outcome, patients who achieved a complete MRD response to blinatumomab within 1 cycle achieved a statistically significantly greater OS (p=0.002) than those who did not respond, as presented in Figure 14. Median OS for 1 cycle complete responders was 38.9 months, compared with 10.5 in non-responders. Similarly, the 18-month OS rate was higher in patients who achieved a complete MRD response to blinatumomab within 1 cycle than those who did not (69% versus 31%).[56]

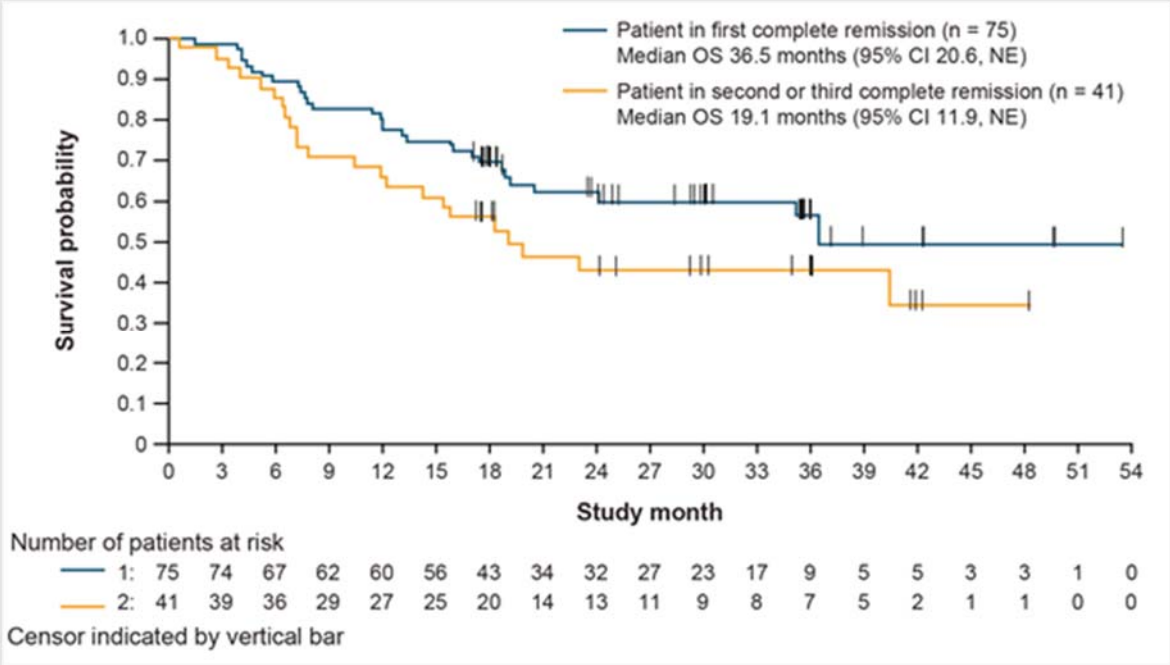
Figure 14. OS in patients with or without a complete MRD response in cycle 1



Abbreviations: OS: Overall survival; MRD: minimal residual disease.
Source: BLAST Key Secondary Analysis CSR Figure 14-4.3.9[56]

Furthermore, relapse history was also associated with median OS; patients who were in their first haematological CR had a longer OS than those who were in their second or third CR (36.5 months versus 19.1 months), as well as a higher 18-month OS rate (69% versus 56%), as presented in Figure 15.[56]

Figure 15. OS in patients in first haematological remission or second/third remission



Abbreviations: OS: Overall survival.
Source: BLAST Key Secondary Analysis CSR Figure 14-4.3.12[56]

Mortality rate within 100 days after allogeneic HSCT

- **Of the 77% of patients who received allogeneic HSCT after treatment with blinatumomab, the 100-day mortality rate was 7%compared to published data >25%.**

The overall patient incidence of HSCT after treatment with blinatumomab was 77.6% (90/116); of these 90 patients, 84.4% were in complete haematological CR at the time of HSCT, with 21.1% being MRD+ and 63.3% MRD− at the end of cycle 1, while 15.6% had haematological relapse prior to HSCT.[56] Published data have shown a 100-day HSCT mortality rate of 28%.

In the FAS, 74 patients received an allogeneic HSCT while in blinatumomab-induced molecular remission. Of these patients, 5 deaths (16.1%) occurred during the 100 days post-HSCT period, resulting in a 100-day mortality rate after allogeneic HSCT of 7% (95% CI: 0.03, 0.15).[56]

Time to haematological response

- **TTHR at 18 months was 55% for patients treated with blinatumomab and censored at HSCT or post-blinatumomab chemotherapy, while median TTHR was not estimable after more than 40 months.**
- **For the uncensored analysis, TTHR at 18 months was 67%, while median TTHR was also not estimable.**

TTHR was measured from the start of treatment with blinatumomab until the patient experienced haematological or extramedullary relapse; patients who died or received HSCT or post-blinatumomab chemotherapy were censored at their last haematological assessment prior to death or post-blinatumomab therapy, whichever occurred first.[56]

As presented in Table 22, the 18-month TTHR censored at HSCT or post-blinatumomab chemotherapy was 55%, while the median TTHR was not estimable. A total of 82.7% of patients were censored as of the data cut-off, and a total of 17.3% had events: 16.4% relapsed and 0.9% had secondary leukaemia. The 18-month KM estimate for TTHR, not censored for HSCT or post-blinatumomab chemotherapy, was 67% (95% CI: 57%, 76%), and the median TTHR was not estimable (95% CI: 24.3, n.e.).[56]

Table 22. TTHR

Outcome	Blinatumomab (N=110)	Blinatumomab censored (N=110)
Events, n (%)	39 (35.5)	19 (17.3)
Relapse	37 (33.6)	18 (16.4)
Secondary leukaemia	1 (0.9)	1 (0.9)
Censors, n (%)	71 (64.5)	91 (82.7)
TTHR (18 months)	0.67	0.55
95% CI	0.57, 0.76	0.34, 0.72
Median (95% CI)	n.e. (24.3, n.e.)	n.e. (7.1, n.e.)

Abbreviations: TTHR: time to haematological relapse; CI: Confidence Interval; n.e.: not estimable.
Source: BLAST Key Secondary Analysis CSR Table 14-4.3.3[56]

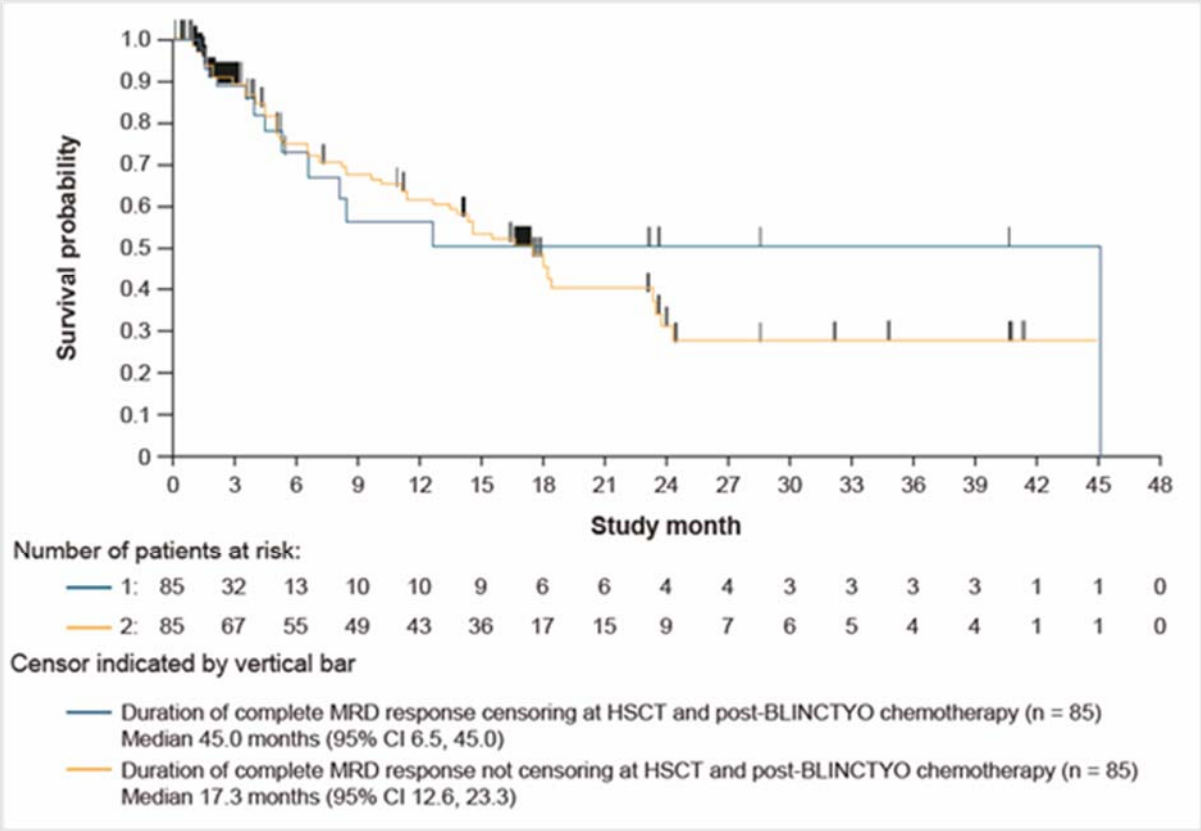
Duration of complete MRD response

- **Blinatumomab offers durable complete MRD responses, with a median duration of 17.3 months, or when censored by HSCT or post-blinatumomab chemotherapy of 45.0 months.**

The median duration of complete MRD response was analysed as the time from onset of MRD negativity until MRD or haematological relapse, or date of last confirmation of negative MRD status. Only the patients with MRD- CR at cycle 1 were included in this analysis, and were analysed both with and without censoring at the time of HSCT or post-blinatumomab chemotherapy.[56]

Blinatumomab offers durable complete MRD responses, with a median duration of 17.3 months (95% CI: 12.6, 23.3) when uncensored and 45.0 months (95% CI: 6.5, 45.0) when censored at HSCT or post-blinatumomab chemotherapy, as presented in Figure 16. The 18-month KM estimates were 46% (95% CI: 33%, 57%) and 51% (95% CI: 28%, 69%), respectively.[56]

Figure 16. Duration of MRD complete response



Abbreviations: MRD: minimal residual disease, HSCT: haematopoietic stem cell transplant.
Source: BLAST Key Secondary Analysis CSR Figure 14-4.5.1[56]

Table 23. Duration of complete MRD response

Outcome	Blinatumomab (N=85)	Blinatumomab censored (N=85)
Events, n (%)	45 (52.9)	16 (18.8)
Censors, n (%)	40 (47.1)	69 (81.2)
Duration of MRD CR (18 months)	0.46	0.51
95% CI	0.33, 0.57	0.28, 0.69

Abbreviations: MRD: minimal residual disease; CR: complete response.
Source: BLAST Key Secondary Analysis CSR Table 14-4.5.1[56]

Effect on MRD level

- **The majority of patients (78%) achieved MRD CR during cycle 1, but of those patients who did not, the majority shifted to a lower MRD status at the end of cycle 1.**

As described in the primary outcome section above, 77.9% of patients achieved MRD CR during cycle 1, with some patients achieving MRD CR as early as 5 days after initiation of treatment. A thorough kinetic analysis of MRD response was not possible, due to the lack of protocol requirement for evaluating MRD response before the completion of cycle 1, combined with the very high rate of complete MRD response by most patients during this first cycle. Nonetheless, it was possible to analyse the MRD response in MRD non-responders from baseline to the end of

cycle 1. A majority of patients [redacted] reported shifts to a lower MRD status at the end of cycle 1, with [redacted] patients with MRD of 10^{-1} shifting to 10^{-5}, [redacted] patients with MRD of 10^{-2} shifting to $\leq 10^{-3}$, and [redacted] patients with MRD of 10^{-3} shifting to $\leq 10^{-4}$. These results demonstrate that even in the minority of patients without a complete MRD response after cycle 1, the majority still improve their MRD status after treatment with blinatumomab. More detail on this outcome is provided in Appendix O.[56]

HRQoL

- **EORTC-QLQ-C30 results suggest that blinatumomab may affect several aspects of HRQoL, such as appetite, constipation, and nausea and vomiting, but most of these showed partial or complete recovery by study end.**
- **EQ-5D results suggest no appreciable change in patient HRQoL during blinatumomab treatment.**

EORTC-QLQ-C30

The EORTC QLQ-C30 is a validated patient reported outcome (PRO) questionnaire comprising multi-item scales and single-item measures rated from 0 to 100 used to assess HRQoL in cancer patients who participate in clinical trials.[62] The questionnaire includes 5 functional scales, 3 multi-item scales, 6 single item symptom scales, and a global health status/quality of life scale. In each of these scales, the patient's quality of life is evaluated on a scale from 1 to 100. Changes of between 5 and 10 points on the EORTC QLQ-C30 scales can be considered clinically meaningful.[63] Change from baseline in the EORTC-QLQ-C30 was analysed using the FAS at each scheduled assessment.[56]

A summary of the maximum changes from baseline to cycles 1 through 4 and to the end of the core study is presented in Table 24. The scales most severely affected by treatment with blinatumomab were appetite loss, constipation, and diarrhoea, and, to a lesser extent, nausea and vomiting. For dyspnoea, constipation, and diarrhoea, scores recovered to baseline at end of core study; for appetite loss and nausea and vomiting, the scores showed partial recovery.

Modest improvements in social functioning and role functioning symptoms were reported during the study; the improvement in social functioning symptom persisted at end of core study. Treatment with blinatumomab provided patients with quality of life improvements in some sub-scales, with modest improvements in the social functioning and role functioning symptoms during the study, which persisted at the end of the core study for social functioning [redacted] [56]

Table 24. Change from baseline in EORTC-QLQ-C30 scales

EORTC-QLQ-C30 Scale	Baseline, mean (SE) (Max=100)	Greatest change from baseline in cycles 1 to 4, mean (SE)/cycle	Change from baseline at end of core study, mean (SE)
Global health status	[redacted]	[redacted]	[redacted]
Physical function	[redacted]	[redacted]	[redacted]
Role functioning	[redacted]	[redacted]	[redacted]
Emotional functioning	[redacted]	[redacted]	[redacted]
Cognitive functioning	[redacted]	[redacted]	[redacted]

Social functioning	[redacted]	[redacted]	[redacted]
Fatigue	[redacted]	[redacted]	[redacted]
Nausea and vomiting	[redacted]	[redacted]	[redacted]
Pain	[redacted]	[redacted]	[redacted]
Dyspnoea	[redacted]	[redacted]	[redacted]
Insomnia	[redacted]	[redacted]	[redacted]
Appetite loss	[redacted]	[redacted]	[redacted]
Constipation	[redacted]	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]	[redacted]
Financial difficulties	[redacted]	[redacted]	[redacted]

Abbreviations: EORTC-QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-C30; SE: Standard Error.
Source: BLAST Key Secondary Analysis CSR Table 11-1[56]

EQ-5D

The EQ-5D is a self-administered PRO which captures 5 dimensions of health: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Change from baseline in the EQ-5D was analysed using the FAS at each scheduled assessment.[56]

A summary of the maximum changes from baseline to cycles 1 through 4 and to the end of the core study is presented in Table 25. Across the 5 dimensions, patients did not experience any appreciable change in quality of life during treatment with blinatumomab.[56]

Table 25. Change from baseline in EQ-5D scales

EQ-5D scale	Baseline, mean (SE)	Greatest change from baseline in cycles 1 to 4, mean (SE)/cycle	Change from baseline at end of core study, mean (SE)
Mobility	1.2 (0.0)	-0.2 (0.1)/C4	0 (0.1)
Self-care	1.1 (0.0)	-0.1 (0.1)/C4	0 (0.0)
Usual activity	1.5 (0.1)	-0.1 (0.1)/C3 + C4	-0.1 (0.1)
Pain/discomfort	1.4 (0.0)	-0.2 (0.2)/C4	-0.1 (0.1)
Anxiety/depression	1.4 (0.1)	-0.2 (0.1)/C2	-0.1 (0.1)

Abbreviations: EQ-5D: EuroQol five dimensions questionnaire; SE: Standard Error.
Source: BLAST Key Secondary Analysis CSR Table 11-2[56]

B.2.6.2 MT103-202

Summary of MT103-202

- Blinatumomab achieved MRD negativity in 80% of BCP-ALL patients in haematological remission.

- Median duration of complete MRD response for patients was 13.0 months, while the median haematological RFS had not been reached after a median follow up time of 1550 days (> 4 years).
- Five of the 9 patients who received HSCT remained in haematological CR for at least 5 years after starting blinatumomab, as did 5 of the 11 patients who did not undergo HSCT, suggesting that long-term disease control can be achieved with blinatumomab with or without subsequent HSCT.

This section provides an overview of the clinical effectiveness results for the MT103-202 pilot study, including the primary outcome, i.e. the proportion of patients who achieved a complete MRD response within 4 cycles of blinatumomab, and the following secondary outcomes: MRD CR after any cycle, TTHR, MRD progression, and MRD relapse after any cycle.[57]

MT103-202 met its primary endpoint, with most patients achieving MRD response (80%; 95% CI: 56.3, 94.3), all of which were achieved in cycle 1. The median duration of complete MRD response for patients was 13.0 months (95% CI: 2.8, not estimable). The median haematological RFS had not been reached after a median follow up time of 1550 days (> 4 years). Ten patients (10/20, 50%) were relapse free after 5 years of follow-up (duration of follow-up ranged from 1816 to 2138 days). The final haematological RFS estimate was 52.6% after 5.9 years. Five of the 9 patients who received HSCT remained in haematological CR for at least 5 years after starting blinatumomab, as did 5 of the 11 patients who did not undergo HSCT, suggesting that long-term disease control can be achieved with blinatumomab with or without subsequent HSCT. These results demonstrate the ability of blinatumomab to achieve MRD negativity in BCP-ALL patients in haematological remission over the long-term, and provide additional support to the findings of BLAST.[57]

B.2.7 *Subgroup analysis*

To determine the impact of biologic predictors of response on study outcomes in BLAST, pre-specified subgroup analyses were defined by a range of baseline variables and were conducted for the primary outcome (MRD CR within 1 cycle), key secondary outcome (haematological RFS rate), other secondary outcomes (OS, TTHR), and HSCT status. An overview of the pre-specified subgroups explored in BLAST are presented in Table 26 below.[56]

Table 26. Overview of pre-specified subgroups in BLAST

Stratification factor	Specific subgroups tested
Age	18–34, 35–54, 55–64, ≥65
Gender	Male, female
Philadelphia status	Philadelphia positive, Philadelphia negative
Patients by t(4;11) translocation and/or MLLAF4+ ALL haematological remission	Yes, No, Unknown
Risk stratification	Standard, Low, Intermediate, High, Very high, Unknown
Relapse history	Patients in 1st CR, Patients in 2nd CR, Patients in 3rd CR

MRD level at baseline by central lab	≥10xE-1 and <10xE0, ≥10xE-2 and <10xE-1, ≥10xE-3 and <10xE-2, <10xE-3, Below LLOQ, Unknown
WBC at first diagnosis	≤30,000/mm^3, >30,000/mm^3, Unknown
Chemoresistance after the first week of chemotherapy	Yes, No, Unknown
Need of salvage therapy for CR	Yes, No, Unknown
Previous anti-tumour radiotherapies	Yes, Unknown
Incidence of neurologic events during cycle 1	Yes, No
Time from diagnosis to start of blinatumomab	≤ 12 months, > 12 months
Time from last treatment to start of blinatumomab	≤ 6 months, > 6 months
Clinical Trial Material	CTM4 only, CTM5 only, CTM4 & CTM5

Abbreviations: ALL: acute lymphoblastic leukaemia; CR: complete remission/complete response; MRD: Minimal residual disease; WBC: white blood cell; LLOQ: lower limit of quantification; CTM4/5: clinical trial material from manufacturing process 4/5.
Source: BLAST Key Secondary Analysis CSR Table 14-4.1.3[56]

A summary of the results for the subgroup analyses is provided in Appendix D. These analyses support the ability of blinatumomab to induce a complete MRD response in patients with MRD+ BCP-ALL in haematological CR, regardless of age, gender, risk stratification, relapse history, or any other subgroup listed in Table 26. For RFS, only relapse history was found to have a significant effect, with a significantly shorter median RFS for patients in their 2nd or 3rd CR (24.6 versus 11.0 months, p=0.0044). OS was only affected by Ph status, with Ph− patients experiencing a significantly longer median OS than Ph+ patients (36.5 versus 7.2 months, p=0.017); however, with only 5 Ph+ patients enrolled in the study, this finding may not be representative of the Ph+ population. As may be expected, median TTHR was significantly shorter in patients in their 2nd or 3rd CR than those in their 1st (19.1 versus 36.5 months, p=0.087).[56]

B.2.8 *Meta-analysis*

No meta-analyses were carried out as only one Phase II study (BLAST) was identified and no other comparator interventions for MRD+ patients are known.

B.2.9 *Indirect and mixed treatment comparisons*

Summary of comparative effectiveness

- Due to the very low incidence of MRD+ BCP-ALL, as well as ethicl and consent issues, it was agreed following discussions with regulatory bodies that BLAST be designed as a single-arm trial based on the successful achievement of MRD negativity in the pilot study.
- As such, a historical comparator study (Study 20120148) was developed to provide a well-

matched cohort of patients treated with SoC, therefore permitting comparison to blinatumomab.

- Study 20120148 included RFS and OS, and was assembled from databases of ALL study groups across Europe, and a direct comparison analysis set was designed *post hoc* to include patients most closely matched to the BLAST population.
  - While this resulted in a well-balanced cohort, limitations of this method include differences in local MRD testing protocols, the timing of MRD assessment, transplant status, and number of prior treatments received.
  - It is noteworthy that in a recent NICE appraisal of blinatumomab for relapsed or refractory (R/R) Ph− BCP-ALL,[68] a similar comparison was made between a historical comparator study and a Phase II trial, and was found to be highly consistent with and validated the comparative results from the pivotal Phase III study.
- Propensity score matching was used to permit comparison of BLAST to the historical comparator study. As the BLAST trial was expected to more closely match the anticipated licensed population than the historical controls, the average treatment effect on the treated (ATT) weighting method was considered most appropriate.
- Blinatumomab was associated with a statistically significant improvement in RFS compared to SoC in patients with MRD+ BCP-ALL, and a consistent numerical improvement in OS compared to SoC:
  - Compared with SoC, blinatumomab reduces the risk of relapse or death by [redacted] and increases the median haematological RFS by [redacted]
  - Compared with SoC, blinatumomab reduces the risk of death by [redacted] and [redacted] the median OS in BCP-ALL patients in MRD+ haematological CR.
- Scenario analyses using an alternative weighting method, the average treatment effect (ATE), are consistent with the ATT weights used to inform estimates of the comparative efficacy.

B.2.9.1 Rationale

Due to the very low incidence of MRD+ BCP-ALL, conducting large randomised clinical studies in this patient population is complex and RCT data are non-existent. In cases where there is no clear standard of care or where currently available therapies have limited evidence of efficacy, it would be inappropriate and unethical to randomise patients to placebo and a single-arm trial is the appropriate choice of trial design. Furthermore, patients and their physicians may not consent to being randomised to placebo or SoC, particularly when blinded studies are not possible. As such, BLAST was designed as a single-arm trial because of the high rate of complete MRD response observed in the pilot trial, the anticipated high rate of complete MRD response in BLAST, and the poor responses to current SoC treatments for BCP-ALL patients in MRD+ haematological CR. In their Scientific Advice provided on December 17th, 2009, the Committee for Medicinal Products for Human Use stated that it could accept data from a single-arm trial if "good quality comparative controls would be available which would well match the patient population in the proposed confirmatory study." The population enrolled in BLAST was highly selected and identifiable, with well-defined baseline characteristics. Thus, a retrospective study of historical data was performed to collect data from patients who represent high quality external controls, matched to the BLAST participants.[65]

Therefore, to better understand the benefit of blinatumomab treatment with respect to RFS and OS among adult MRD+ BCP-ALL patients relative to historical controls, a propensity score analysis was applied to quantitatively evaluate these endpoints. In this section, an overview of the historical comparator study is provided, and results for the propensity score analysis of BLAST and the historical comparator study are presented.[65]

B.2.9.2 Historical comparator study (Study 20120148)

This retrospective, historical comparator study evaluated clinical outcomes, including RFS and OS, in BCP-ALL Ph− adult patients who have received SoC treatment per national treatment of study group protocols, achieved a haematological CR, and subsequently had persistent or relapsed MRD. Data were collected from study groups across Europe and Russia, and clinical data for 287 patients with similar baseline characteristics to those enrolled in BLAST were included in the analysis.[58]

A summary of the design of the historical comparator study is presented in Table 27. The inclusion criteria were chosen such that the patients enrolled in the study were matched as closely as possible to the characteristics of those in the BLAST trial, including history of ALL treatment (response to first therapy and number of prior relapses), relapse status, and disease follow-up after detection of MRD.[58]

Table 27. Summary of design for historical comparator study (Study 20120148)

Study description	Retrospective non-interventional cohort study of historical treatment and outcome data from 2000 to 2017 for 287 adult patients
Patient population eligibility	Patients with BCP-ALL Ph− with haematological CR (defined as less than 5% blasts in bone marrow after at least 3 intensive chemotherapy blocks, and who met the following criteria: • Detection of MRD (molecular failure or molecular relapse) at a level of ≥10^{-4} by PCR or ≥10^{-3} by flow cytometry at a reference lab • Age 15+ at time of initial diagnosis of ALL. For patients 15-17 years of age at diagnosis, patients were not allowed to be enrolled in a paediatric trial • Initial diagnosis of ALL in the year 2000 or later • History of ALL treatment (including response to first therapy, number of prior relapses) is available • Relapse status and disease follow-up after time point of MRD detection is available
Exclusion criteria	• Patients with extramedullary disease at timepoint of MRD detection • Use of blinatumomab within 18 months of MRD detection • Allogeneic HSCT prior to MRD detection at required level
Primary endpoint	• Haematological RFS, defined as the time from the baseline MRD detection date until haematologic relapse or death due to any cause
Key secondary endpoints	• OS, defined as the time from the baseline MRD detection date until death • Mortality rate (proportion) in patients who received an allogeneic HSCT after MRD detection, assessed at 100 days following allogeneic HSCT, as well as later timepoints (3, 6, 9 and 12 months, and 6-monthly intervals until 36 months after allogeneic HSCT).

Abbreviations: BCP: B-Cell Positive; ALL: acute lymphoblastic leukaemia; MRD: minimal residual disease; HSCT: haematopoietic stem cell transplant; RFS: relapse-free survival; OS: Overall survival.
Source: Historical Comparator Study (20120148) CSR[58]

B.2.9.3 Summary of patient baseline characteristics in BLAST and the historical comparator study

The study population for the historical comparator study was assembled from databases of ALL study groups in Europe (Czech Republic, France, Germany, United Kingdom, Italy, Poland, and Spain), and in Russia, which included MRD testing in their protocols. However, all but Russia contributed data to the direct comparison analysis set (DCAS, which was designed *post hoc* to include patients most closely matched to the BLAST population). For Russian patients, MRD test results were qualitative only, with results above 10^{-4}, but the actual MRD level was not quantified so it was not possible to assess whether patients qualified for the primary analysis sets. For Poland, Spain and one site from Italy, MRD levels were assessed by flow cytometry rather than PCR, therefore none of their patients were included in the primary analysis set, but there were patients from these countries in the DCAS. The primary contributors to the DCAS were Germany (38.5%), Italy (25.8%), France (13.7%) and Poland (13.7%). These study groups were selected based on recognition of the 'state of the art' ALL care and treatment that they provide, and because they included MRD testing in their protocols, thus being relevant and generalisable to the UK, in which the similar UKALL14 protocol is used.[45] Similarly, standardised treatment protocols developed as part of the European Working Group for Acute Lymphocytic Leukaemia (EWALL) collaboration should also allow the comparison of other European study centres to the UK. Indeed, these European and Russian sites also contributed patients to the BLAST study; historical patients may have been affected by a similar bias, given that the majority were enrolled in trials as well. On balance, it was judged that historical data from these study groups on their patients' experience would provide a reasonable population to provide a frame of reference for evaluating the experience of blinatumomab-treated ALL patients.

Limitations

However, there are limitations in the method of historical comparison used. Firstly, each country or study group followed their national or local protocol, which specified when to conduct MRD assessments. Therefore, the timing of MRD assessment following initial diagnosis varied between countries or study groups and the overall estimate of timing of the baseline MRD assessment reflects the average time as defined in these treatment protocols. This average was driven mainly by the countries or study groups that contributed the largest number of patients (i.e. Germany, a site in Italy and France). Given that baseline MRD assessment was the time point from which RFS and OS were assessed in the PAS, this may have led to bias in the duration of survival.

Secondly, a potential source of bias comes from differences in the transplantation status of patients, as transplanted patients are likely to be younger than non-transplanted patients, and may have a better general health status or other unmeasured characteristics (e.g., finding a suitable donor) that are systematically different from patients who did not undergo transplantation. Thus, any observed differences in the standard KM curves for RFS and OS between transplanted and non-transplanted patients should be interpreted with caution as these differences may not be because of transplant only.

Finally, the number of prior treatments received could also potentially bias results. Patients from all countries were required to have been treated with at least 3 intensive chemotherapy blocks before qualifying to enter the study, except for patients from the UK; these patients were eligible to enter the study after being treated with only two blocks. However, this exception was made

because in terms of intensity, the first 2 blocks of treatment in the UK protocol were considered to be comparable to 3 blocks of other study groups.

While it is important to note these limitations, the historical comparator study was specifically designed to be comparable to the blinatumomab clinical studies, using strict eligibility criteria, and, as previously described, the DCAS was designed *post hoc* to include patients most closely matched to the BLAST population. Nevertheless, scenario analysis were conducted to assess the potential impact of the comparative evidence on cost-effectiveness as discussed in Section B.8.

Baseline characteristics

A summary of the baseline characteristics of patients enrolled in BLAST (from the FAS) and the historical comparator study (from the DCAS) is presented in Table 28.

Although the inclusion criteria in the historical comparator study allowed the identification of well-matched patients, there were some differences in baseline characteristics between Study 20120148 and BLAST. Median patient age in the DCAS population was lower than in the BLAST FAS population (33 versus 45 years), and only one patient was 65 years of age or older, compared with 15 patients in BLAST. The DCAS did not include any patient in CR2 or CR3. However, 41 patients were in their second or third CR in the BLAST FAS; therefore, only patients in CR1 were included in the primary analysis.[65]

Table 28. Patient baseline characteristics in BLAST and the historical comparator study

Demographic	BLAST	Historical comparator
	FAS[a] n=116	DCAS[b] n=182
Male sex, n (%)	68 (59)	102 (56)
Median age (range), years	45.0 (18–76)	33.0 (18–65)
Age, n (%) ≥18 to <35 years	36 (31.0)	98 (53.9)
≥35 to <55 years	41 (35.3)	56 (31.0)
≥55 to <65 years	24 (20.7)	27 (14.8)
≥65 years	15 (12.9)	1 (0.6)
Median time from prior treatment (range), months	[redacted]	NA
Relapse history, n (%) First CR	75 (64.7)	182 (100)
Second CR	39 (33.6)	NA
Third CR	2 (1.7)	NA
Baseline MRD levels, n (%) $\geq 10^{-1} < 1$	[redacted]	13 (7.1)
$\geq 10^{-2} < 10^{-1}$	[redacted]	65 (35.7)
$\geq 10^{-3} < 10^{-2}$	[redacted]	104 (57.1)
$< 10^{-3}$	[redacted]	0 (0)
Below LLQ	[redacted]	NA

Unknown	[redacted]	0 (0)

Abbreviations: [a]All patients receiving a BLINCYTO infusion; [b]Patients ≥18 years old with MRD load ≥1 × 10−3 detected by FC or PCR in CR1, time to haematological relapse >14 days after MRD diagnosis.
CR: complete remission; DCAS: direct comparison analysis set; FAS: full analysis set; LLQ: lower limit of quantification; NA: not applicable.
Source: BLAST Key Secondary Analysis CSR,[56] Historical Comparator Study (20120148) CSR[58]

B.2.9.4 Propensity score analysis

Because of these differences in baseline characteristics between the historical comparator study and BLAST populations, propensity scoring (PS) was performed to balance the baseline covariates between the groups to allow a more valid statistical comparison of RFS and OS between the two trials; the inverse probability of treatment weighting (IPTW) approach for propensity score adjustment was used for this analysis.[65] This approach attempts to mimic the effect of randomisation by creating a balance between treated and untreated patients with respect to important baseline covariates that determine both the propensity for a patient to be treated with the treatment under evaluation (in this case, blinatumomab) and a patient's prognosis. Further details of the propensity score model are provided below with further key details summarised in Appendix L[65].

Analysis sets

The Primary Analysis Set for the PS analysis included patients who adhered to the following criteria:

BLAST (MT103-203) criteria:

- Received any infusion of the investigational drug, blinatumomab
- Philadelphia negative B–precursor ALL in complete haematological remission defined as less than 5% blasts in bone marrow after at least three intensive chemotherapy blocks
- MRD+ at a level of ≥1 x 10^{-3} (PCR only in BLAST) but otherwise in complete haematological remission
- At least 18 years old at the MRD baseline date
- In their first haematological remission (CR1 only)

Historical comparator study (20120148) criteria:

- Ph-negative B-precursor ALL in complete haematological remission
- MRD+ at a level of ≥1 x 10^{-3} regardless of detection method
- At least 18 years old at the MRD baseline date
- Time to relapse greater than 14 days from the date of MRD detection (see explanation below)
- In their first haematological remission (CR1 only)

Propensity Score model development

Candidate variable main effects and all two-way interaction terms were entered into a logistic regression model with blinatumomab treatment as the binary response. A stepwise variable selection algorithm was run whereby $p<0.30$ was used as the threshold for entering and keeping covariates in the model.

Candidate propensity score model covariates included the following:

- Age at primary diagnosis
- Sex
- Country
- Presence and type of any cytogenetic and molecular aberrations
- Time from primary diagnosis to MRD baseline date (months)
- Baseline MRD level
- WBCs at diagnosis
- Type of prior chemotherapy

Baseline MRD was recoded into an ordinal variable and was treated as a continuous covariate in the model. A propensity score model was fit for each analysis set separately.

The PS-weighted RFS and OS analyses was performed using a Cox proportional hazards model with each patient's treatment status as an independent factor. An additional analysis including a time-dependent covariate for HSCT was conducted to account for differences between transplant rates observed between BLAST and the historical cohort and better isolate the blinatumomab treatment effect not affected by use of transplant. Robust variance estimation was applied to all models using the COVSANDWICH option in PROC PHREG in SAS.

The estimated survival probabilities from the Cox survival model before and after the use of PS weights were plotted for comparisons. Survival rates and their 95% CIs were estimated based on the Cox survival model and Kaplan Meier curves with median and 95% CIs were estimated.

Propensity Score Adjustment Method

The IPTW approach for propensity score adjustment was used for this analysis where different weights can be applied depending on the objective of the analysis. In this analysis, two sets of weighting approaches have been explored: the average treatment effect on the treated (ATT) and the average treatment effect (ATE). It should be noted that the pre-specified protocol for the propensity score analysis provided for ATE analyses as the primary analysis and these were the data presented to the regulator for licensing. ATE analyses adjust (weight) both the treated and untreated populations by assuming that they are drawn from one homogenous population, requiring strong assumptions of ignorability and overlap between the studies.[66] In contrast, identifying the ATT, which adjusts (weights) the control population only, requires ignorability in mean but only for the outcome without treatment and a weaker version of overlap.[66] Given the orphan nature of this disease, and the consequent small sample size in the historical comparator study, a matching analysis could not be undertaken to help further address the issue of comparability.

In considering which of the ATE and ATT populations are most generalisable to the present appraisal, it should be noted that the trial population in the BLAST study represents the prospectively selected anticipated licensed population, whereas that in the historical comparator represents a retrospectively identified population; given this, the population of interest for this appraisal is likely to be the treated (ATT) population rather than a weighted mixture of the treated and untreated (ATE) population (which reduces the weight given to some treated patients in order to more closely match the population in the historical comparator study).

Given this, the ATT weighting was the preferred approach, as results based on ATT weight can be generalised to the population of patients in BLAST rather than the combined populations of the BLAST and historical control studies. This is appropriate as the historical control study was designed *a priori* to match patients in BLAST, and the use of propensity score analysis was conducted only to control for residual confounding after matching. Although ATT weights were used for base-case analyses, a scenario analysis was conducted using ATE weights; results of the ATE analysis are presented in Appendix L and the impact of using these weights are explored in a scenario analysis in the economic evaluation (Section B.3.8.3).

Results (ATT weights)

Balance in baseline covariates using ATT weights

The balance between treatment groups with respect to the 8 covariates, as well as a continuous measure of WBC at diagnosis (log transformed), was assessed both before and after adjustment and is summarised in Table 29. After adjustment, none of the p-values were significant and [redacted] covariates had standard differences less than [redacted]. Three covariates with standard differences greater than [redacted] were: age, time from primary diagnosis to baseline and WBC at diagnosis.

Table 29. Covariate Balance Before and After Propensity Score Adjustments using ATT weights

Characteristic	Unweighted				Stabilised IPTW			
Mean (SD)/n (%)	Control [redacted]	Blinatumomab [redacted]	Standard Difference	P-value	Control [redacted]	Blinatumomab [redacted]	Standard Difference	P-value
Age at primary diagnosis (years)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gender (Female)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Country (Not Germany)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
MRD at Baseline (recoded)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Time from diagnosis to baseline (months)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
WBC at diagnosis (>30,000/mm^3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
WBC at diagnosis (continuous, log10)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
T411mll4 mutation (Yes)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Prior chemotherapy (GMALL)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: IPTW: inverse probability of treatment weighting; SD: standard deviation; MRD: minimal residual disease; WBC: white blood cell; GMALL: German Multicentre ALL Working Group.
Source: Propensity score analysis report[65]

RFS

- **Compared with SoC, blinatumomab reduces the risk of haematological relapse or death [redacted] and more than quintuples the median RFS in MRD+ BCP-ALL patients in haematological CR.**

In the analysis without censoring for HSCT, there was a [redacted] reduction in the risk of relapse or death ([redacted]) associated with blinatumomab versus SoC. The median haematological RFS was [redacted] with blinatumomab compared with SoC chemotherapy ([redacted]), and the 18-month RFS rate with blinatumomab was [redacted] that with SoC chemotherapy ([redacted] The K-M curves presented in Figure 17 demonstrate a clear separation in RFS over time between blinatumomab and SoC chemotherapy. This result is robust and is supported across both ATT and ATE weighting methods (see Appendix L for ATE results). When censoring for HSCT, blinatumomab resulted in statistically significantly longer RFS, with a [redacted] reduction in the risk of relapse or death ([redacted]). The similarity of the censored and uncensored analyses suggests that blinatumomab provides improved outcomes independent of transplant.[65]

Figure 17. RFS in BLAST versus SoC chemotherapy using ATT weights

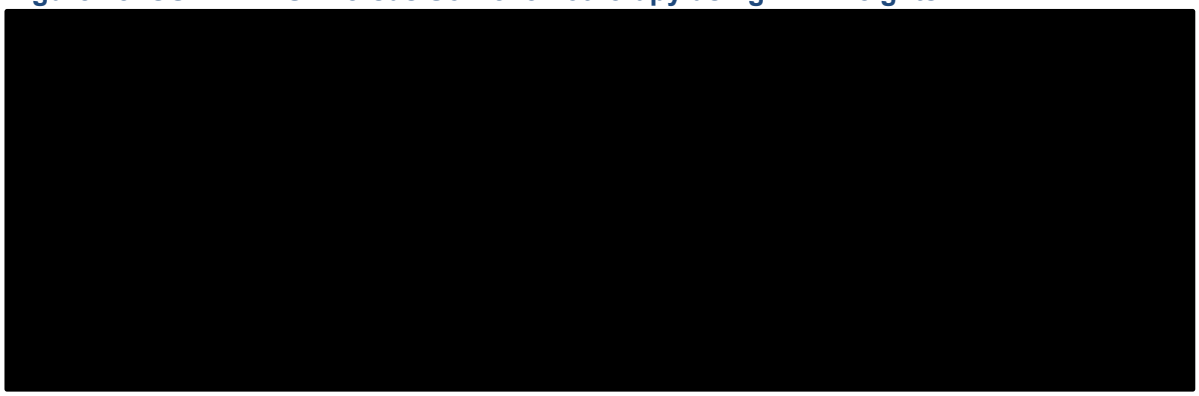
Abbreviations: RFS: relapse-free survival; SoC: Standard of Care; ATT: average treatment effect for the treated patients.
Source: Propensity score analysis report[65]

OS

- **Compared with SoC, blinatumomab reduces the risk of death by [redacted] the median OS in BCP-ALL patients in MRD+ haematological CR.**

Without censoring for HSCT, as in the RFS results, there was a [redacted] reduction in the risk of death [redacted] associated with blinatumomab versus SoC. Median OS was [redacted] after more than 40 months of follow-up, compared to a median of [redacted] for SoC chemotherapy. Furthermore, the 18-month OS was [redacted] greater for blinatumomab compared to SoC chemotherapy ([redacted]). The K-M curves presented in Figure 18 demonstrate similar separation between the survival curves for both ATT and ATE weighting methods (see Appendix L for ATE results). When censoring for HSCT, blinatumomab resulted in statistically significantly longer OS, with a [redacted] reduction in the risk of death compared with SoC chemotherapy ([redacted]). The similarity of the censored and uncensored analyses suggests that blinatumomab provides improved outcomes independent of transplant.[65]

Figure 18. OS in BLAST versus SoC chemotherapy using ATT weights



Abbreviations: OS: Overall survival; SoC: Standard of Care; ATT: average treatment effect for the treated patients.
Source: Propensity score analysis report[65]

B.2.9.5 Uncertainties in the indirect and mixed treatment comparisons

The key uncertainty is the use of an historical control as there are inherent limitations to this analysis. Notably the timing and method of MRD assessment varied by study group and the average time is driven by the largest countries or study groups—baseline MRD assessment was the time point from which RFS and OS were assessed in the PAS, therefore this may have led to bias. The study is also vulnerable to transplantation status being a confounding factor as transplanted patients may be systematically different, being younger, having a better general health status or, crucially, other unmeasured characteristics (e.g., finding a suitable donor). As a result, any differences between transplanted and untransplanted patients may not be because of transplant only. It was also noted that the number of prior treatments received, and other sources of heterogeneity arising from each study group following its own protocol, could also potentially bias the interpretation of any results.

Furhtermore, as noted in Section B.2.9.4, the choice of ATE or ATT analysis is a further point of uncertainty in the comparative analysis presented here. ATE was pre-specified in the analysis plan, however, as the BLAST study population is expected to reflect the anticipated licensed population more closely than the historical comparator study, and therefore the decision problem for this appraisal, it was considered preferable to present the ATT analyses as these inform the economic model inputs. It should be noted that overall ATE and ATT analyses were broadly consistent—finding that blinatumomab was an effective treatment. The ATE analysis is presented as a scenario analysis in the economic model to further address this uncertainty.

Whilst acknowledging the limitations and uncertainties, the historical comparator study was specifically designed to include patients most closely matched to the BLAST population and is anticipated to prove a robust basis for licensing, as well as for informing the economic analysis required in this appraisal.

B.2.10 *Adverse reactions*

Summary of Adverse Reactions

- Blinatumomab has a highly tolerable safety profile; the alternative SoC chemotherapy is associated with substantial adverse effects.

- No new safety signals were observed in patients pooled from the MRD studies, beyond the existing safety profile of blinatumomab.

- Neurological events occurred in 72% of patients receiving blinatumomab, however, more than two thirds of the neurological adverse events associated with blinatumomab were mild to moderate and decreased over time.
- The results of the safety analysis in adult MRD+ BCP-ALL patients reflect the known safety profile of blinatumomab.
- The incidence of adverse events of interest in this patient population, including neurological toxicities, cytokine release syndrome, and medication errors, did not suggest any new risks from blinatumomab beyond those already identified.

The safety and tolerability data presented below are derived from pooled data from all patients who received any infusion of blinatumomab in BLAST or MT103-202, and comparisons have been drawn against the known safety profile of blinatumomab in adult relapsed or refractory Ph− BCP-ALL (pooled data from MT103-206, MT103-211, and TOWER).[67-69]

B.2.10.1 Safety profile of blinatumomab in MRD+ BCP-ALL

The MRD+ BCP-ALL safety analysis included 137 patients: 116 patients from BLAST and 21 patients from MT103-202. Table 30 presents the incidence of treatment-emergent AEs and those considered related to blinatumomab (treatment-related AEs).[70]

Table 30. Incidence of AEs in MRD+ BCP-ALL

Event	Treatment-emergent AEs (n=137)	Treatment-related AEs (n=137)
All AEs, n (%)	137 (100.0)	133 (97.1)
Serious	83 (60.6)	69 (50.4)
Grade ≥3	88 (64.2)	73 (53.3)
Grade ≥4	39 (28.5)	32 (23.4)
Fatal[a]	2 (1.5)	1 (0.7)
Leading to permanent discontinuation of BLINCYTO	23 (16.8)	16 (11.7)
Serious	17 (12.4)	13 (9.5)
Grade ≥3	18 (13.1)	13 (9.5)
Grade ≥4	6 (4.4)	4 (2.9)
Fatal[a]	2 (1.5)	1 (0.7)
Leading to interruption of BLINCYTO	39 (28.5)	35 (25.5)
Serious	29 (21.2)	26 (19.0)
Grade ≥3	22 (16.1)	20 (14.6)
Grade ≥4	8 (5.8)	7 (5.1)
Fatal[a]	0 (0.0)	0 (0.0)

Footnotes: [a]Fatal events that occurred within 30 days of last blinatumomab treatment.
Abbreviations: AE: Adverse Event; BCP: B-Cell Positive; ALL: acute lymphoblastic leukaemia; MRD: minimal residual disease.

Source: Blinatumomab Clinical Overview[70]

All patients experienced at least one treatment-emergent AE; the most common (occurring in ≥20% of patients) were pyrexia (90.5%), headache (39.4%), tremor (29.2%), chills (28.5%), fatigue (26.3%), nausea (23.4%), vomiting (21.2%), hypokalaemia (20.4%), and diarrhoea (20.4%). The rate of treatment-emergent grade ≥3 AEs in MRD+ BCP-ALL patients was lower than that reported in relapsed or refractory Ph− BCP-ALL (64.2% versus 83.9%, respectively). Treatment-emergent serious AEs were reported in 60.6% of MRD+ BCP-ALL patients, a rate that is consistent with the adult relapsed or refractory Ph− BCP-ALL population. The most common (occurring in ≥5% of patients) were pyrexia (12.4%) and tremor (5.8%).[70]

Treatment-related AEs were reported in 97.1% of MRD+ BCP-ALL patients, a higher rate than in the adult relapsed or refractory Ph− BCP-ALL population (84.7%). The most common (>20% patients) were pyrexia (86.1%), headache (27.7%), tremor (27.0%), chills (26.3%), and fatigue (21.2%). Half (50.4%) of all MRD+ BCP-ALL patients were reported to have experienced a treatment-related serious AE, compared with 32.6% of adult relapsed or refractory Ph− BCP-ALL patients. Although the overall rate of treatment-related serious AEs was 50.4%, rates by preferred term were low; only pyrexia (12.4%) and tremor (5.8%) were reported in ≥5% patients. The higher incidence of treatment-related AEs in the MRD+ BCP-ALL population compared with the relapsed or refractory Ph− BCP-ALL population is likely to have been driven by the rate of grade ≤2 AEs, as the rates of grade ≥3 and grade ≥4 AEs were comparable across the two populations (53.3% and 23.4% versus 54.9% and 23.1%, respectively). [70]

There were two deaths due to AEs that occurred within 30 days of the last blinatumomab treatment: one was a fatal infection (atypical pneumonia) and considered related to blinatumomab; the second was a subdural haemorrhage and was not considered to be treatment-related.[70]

Treatment interruptions due to treatment-emergent AEs were required in 28.5% of patients in the MRD+ BCP-ALL population, mainly due to neurological events and flu-like symptoms associated with T-cell activation. Treatment interruptions due to treatment-related AEs were required in 25.5% of patients. The rate of treatment interruptions in the MRD+ BCP-ALL population was consistent with the relapsed or refractory Ph− BCP-ALL population; even with treatment interruptions, 78% of MRD+ BCP-ALL patients achieved MRD negativity with one cycle of blinatumomab treatment (pooled data from BLAST and MT103-202). Twenty-three patients (16.8%) had AEs leading to permanent discontinuation of blinatumomab; the AEs were reported as treatment-related in 16 patients (11.7%).[70]

Table 31 summarises the incidence of treatment-emergent events of interest (EOIs) in MRD+ BCP-ALL patients. The rate of any-grade EOIs was consistent with the adult relapsed or refractory Ph− BCP-ALL population; the rate of grade ≥3 or ≥4 EOIs was lower in the MRD+ BCP-ALL population (56.9% and 27.7% versus 75.8% and 41.9%, respectively).[70]

Table 31. Incidence of treatment-emergent EOIs in MRD+ BCP-ALL

EOIs	All treatment-emergent EOIs	EOIs	All treatment-emergent EOIs
All treatment-emergent EOIs	134 (97.8)	**Medication errors[a]**	6 (4.4)
Serious	74 (54.0)	**Serious**	6 (4.4)

EOIs	All treatment-emergent EOIs	EOIs	All treatment-emergent EOIs
Grade ≥3	78 (56.9)	Grade ≥3	0
Grade ≥4	38 (27.7)	Grade ≥4	0
Fatal	1 (0.7)	Neutropenia and febrile neutropaenia[a]	22 (16.1)
Neurological events[a]	98 (71.5)	Serious	7 (5.1)
Serious	31 (22.6)	Grade ≥3	22 (16.1)
Grade ≥3	22 (16.1)	Grade ≥4	17 (12.4)
Grade ≥4	3 (2.2)	Decreased immunoglobulins[a]	25 (18.2)
Infections	64 (46.7)	Serious	0 (0.0)
Serious	18 (13.1)	Grade ≥3	7 (5.1)
Grade ≥3	16 (11.7)	Grade ≥4	0
Grade ≥4	4 (2.9)	Capillary leak syndrome[a]	1 (0.7)
Fatal	1 (0.7)	Serious	0
Cytokine release syndrome[a]	4 (2.9)	Grade ≥3	0
Serious	2 (1.5)	Grade ≥4	0
Grade ≥3	2 (1.5)	Elevated liver enzymes[a]	17 (12.4)
Grade ≥4	0	Serious	5 (3.6)
Leukoencephalopathy[a]	1 (0.7)	Grade ≥3	11 (8.0)
Serious	1 (0.7)	Grade ≥4	6 (4.4)
Grade ≥3	0	Lymphopaenia[a]	9 (6.6)
Grade ≥4	0	Serious	6 (4.4)
Infusion reaction[a]	124 (90.5)	Grade ≥3	9 (6.6)
Serious	19 (13.9)	Grade ≥4	8 (5.8)
Grade ≥3	14 (10.2)	Pancreatitis[a]	1 (0.7)
Grade ≥4	1 (0.7)	Serious	0
Embolic and thrombotic events[a]	7 (5.1)	Grade ≥3	0
Serious	4 (2.9)	Grade ≥4	0
Grade ≥3	5 (3.6)	Tumour lysis syndrome[a]	0
Grade ≥4	2 (1.5)	Serious	0
		Grade ≥3	0
		Grade ≥4	0

Footnotes: [a]No fatal events were identified in this category
Abbreviations: AE: Adverse Event; EOI: Event of Interest.
Source: Blinatumomab Clinical Overview[70]

Neurological AEs were experienced by 71.5% of MRD+ BCP-ALL patients, a rate similar to that observed in the adult relapsed or refractory Ph− BCP-ALL population (66.9%). Most neurological AEs were mild to moderate; the most common (occurring in ≥10% patients) were headache (39.4%), tremor (29.2%), insomnia (16.1%), aphasia (11.7%), and dizziness (10.2%). Twenty-two patients (16.1%) experienced a neurological AE of at least grade 3; no fatal neurological AEs were reported. The median duration of neurological events was 10 days (95% CI: 6.0, 15.0). Analysis of safety data from BLAST has demonstrated that most treatment-related neurological AEs were mild to moderate; rates of grade ≥3 and grade ≥4 AEs were 12.1% and 2.6%. In BLAST, the rate of neurological AEs decreased with each treatment cycle: from 47% in cycle 1 to 15% in cycle 4, and for grade 3 or higher neurological AEs from 10% in cycle 1 to 0% in cycles 3 and 4.[70]

Four MRD+ BCP-ALL patients (2.9%) were identified as having cytokine release syndrome. This rate was considerably lower than that reported in the adult relapsed or refractory Ph− BCP-ALL population (14.2%) and may be a result of the lower disease burden in this population in haematological CR. Two patients (1.5%) experienced grade 3 cytokine release syndrome; no grade 4 or 5 events were reported. Treatment with blinatumomab was interrupted in one patient because of cytokine release syndrome.[70]

Medication errors (coded as overdose or accidental overdose) occurred in six patients (4.4%) in the MRD+ BCP-ALL population, a rate comparable to the adult relapsed or refractory Ph− BCP-ALL population (3.8%). All medication errors were reported as serious, in accordance with protocol guidance.[70]

B.2.10.2 Summary of safety

The results of the safety analysis in adult MRD+ BCP-ALL patients reflect the known safety profile of blinatumomab. The incidence of EOIs in this patient population, including neurological toxicities, cytokine release syndrome, and medication errors, did not suggest any new risks from blinatumomab beyond those already identified. Indeed, more than two thirds of the neurological adverse events associated with blinatumomab were mild to moderate and decreased over time. The US and European approval of blinatumomab for the treatment of relapsed or refractory Ph− BCP-ALL is contingent on a risk management plan to inform providers and patients (and their caregivers) of the serious risk of neurological toxicities, cytokine release syndrome, and medication errors. The results of the safety analysis in adult MRD+ BCP-ALL patients from BLAST and MT103-202 reflect the known safety profile of blinatumomab. The incidence of EOIs in this patient population, including neurological toxicities, cytokine release syndrome, and medication errors, did not suggest any new risks from blinatumomab beyond those already identified.[70]

B.2.11 *Ongoing studies*

No ongoing studies are expected to provide additional evidence for blinatumomab in MRD+ BCP-ALL in the next 12 months.

B.2.12 *Innovation*

With current SoC chemotherapies, patients with MRD+ BCP-ALL in haematological remission have poor leukaemia-free survival and are less likely to receive successful allogeneic HSCT than MRD negative patients.[48, 71] There are currently no treatment options specifically indicated for

patients in this high-risk population and current chemotherapy regimens used are not effective; therefore, there is considerable unmet need for MRD+ patients. Blinatumomab is the first therapy specifically licensed in this indication, and represents a paradigm shift in how MRD+ patients are managed. As the only therapy indicated specifically for the treatment of adults with MRD+ BCP-ALL in haematological CR, blinatumomab is expected by clinical experts to become the standard of care for this population.[56]

Blinatumomab is a novel single-agent bispecific T-cell engaging immunotherapy with a first-in-class mechanism of action that harnesses the body's own immune system to recognise and eliminate malignant cancer cells (see Section B.1.2). Blinatumomab's innovative mechanism of action facilitates transient connection of malignant cells with T cells, thereby inducing T-cell-mediated killing of the bound malignant cell. By bringing T cells into close proximity with tumour cells much more frequently than without blinatumomab, the surveillance and cytotoxic abilities of the patient's own T cells are greatly increased. This innovative and novel mechanism of action provides clinicians with an alternative treatment option to conventional chemotherapies.[56]

The innovative nature of blinatumomab was demonstrated by being the first bispecific antibody construct to be approved by the FDA, approved 5 months ahead of schedule after receiving "breakthrough therapy" designation in June 2014. Blinatumomab is approved in the EU and US for the treatment of adults with relapsed or refractory Ph− BCP-ALL, and additionally for paediatric patients with relapsed or refractory Ph− BCP-ALL in the US.[72, 73]

With this innovative mechanism of action, blinatumomab can induce a complete MRD response in up to 80% of adult MRD+ BCP-ALL patients in haematological CR. Inducing a complete MRD response in these patients reduces the risk of subsequent relapse, reduces the risk of HSCT failure and improves OS, independent of HSCT. Compared with SoC chemotherapies, blinatumomab reduces the risk of haematological relapse or death [redacted] and [redacted] the median RFS in MRD+ BCP-ALL patients in haematological CR, independently of HSCT (Section B.2.9.4). Compared with SoC, blinatumomab reduces the risk of death by [redacted] and [redacted] the median OS in MRD+ BCP-ALL patients in haematological CR (Section B.2.9.4).

Blinatumomab is also be associated with a number of benefits that may not be captured within the NICE incremental cost-utility framework. For example, blinatumomab is an effective therapy that can be administered in the outpatient setting, and therefore has the potential to reduce duration of hospitalisation compared with current SoC chemotherapy regimens, which require specialist nurses, limited hospital administration facilities and the high costs associated with the management of complications. This may also increase valuable time at home for ALL patients. Furthermore, by achieving sustained relapse-free survival, blinatumomab is expected to lead to an increase in patients being cured of the disease, and is also expected to lead to an increase in patients experiencing positive post-HSCT outcomes; in addition to the important survival benefit this represents, sustained RFS may lead to greater productivity.

Based on the above, the introduction of blinatumomab as a highly innovative and well-tolerated therapy with demonstrable efficacy to achieve MRD negativity after SoC chemotherapy represents a paradigm shift in the management of patients with MRD+ BCP-ALL in haematological CR. These patients currently have no licensed, targeted treatment options available to them and blinatumomab has the potential to help address the considerable unmet medical need for these patients.

B.2.13 *Interpretation of clinical effectiveness and safety evidence*

B.2.13.1 Summary of the principle findings of the clinical evidence base

Clinical evidence

Blinatumomab is the first and only therapy specifically indicated for adults with BCP-ALL in haematological remission that can be used to achieve MRD negativity after SoC therapy.

The confirmatory Phase II BLAST trial in 116 patients supports the findings of MT103-202, and demonstrates the ability of blinatumomab to induce a complete MRD response in the clear majority of patients in haematological CR, irrespective of baseline covariates (Section B.2.6.1). Furthermore, when compared to SoC chemotherapy using a purpose-designed, well-matched historical comparator study, patients treated with blinatumomab demonstrate statistically significantly longer RFS, with a [redacted] reduction in the risk of relapse or death and [redacted] median RFS (Section B.2.9.4). Similarly, compared to SoC, patients treated with blinatumomab also demonstrate statistically significantly [redacted] OS, with a [redacted] reduction in the overall risk of death, while median OS was still [redacted] even after 40 months' follow-up (Section B.2.9.4). Those patients who demonstrate a complete MRD response when treated with blinatumomab also demonstrate a significantly longer RFS and OS than non-responders. Measures of response in ALL were objectively defined by laboratory results, measured by a central laboratory, rendering them less prone to bias from the physician or patient than may be the case in other cancer indications.[56]

None of the subgroups tested were determined to have a statistically significant effect on the primary outcome of complete MRD response, demonstrating the robust treatment effect provided by blinatumomab, even across older patients and those with a high level of MRD; both populations which currently have limited treatment options. Only relapse history was considered to have a statistically significant effect on RFS (P = 0.0044), and also on TTHR (P = 0.0031).

Health-related quality of life evidence

Patient HRQoL during treated with blinatumomab was assessed by the EORTC-QLQ-C30 and EQ-5D scales. While patients treated with blinatumomab experience increased rates of appetite loss, constipation, diarrhoea, and, to a lesser extent, nausea and vomiting, and dyspnoea, the majority of these showed full or partial recovery to baseline levels at the end of the core study.

Furthermore, across all dimensions of the EQ-5D, patients experienced no reduction in HRQoL, while in the EORTC-QLQ-C30 scales of social functioning and role functioning patients experienced modest improvements, which persisted at the end of the core study.

A key benefit of blinatumomab is its position in the treatment pathway for MRD+ BCP-ALL patients in haematological remission. MRD positive status is the key predictor of disease relapse, and there are currently no approved therapies for these patients. The psychological impact on such patients, in being told that there are no available treatments, may be considerable. Blinatumomab, as the first therapy specifically indicated for patients with MRD+ BCP-ALL, represents a paradigm shift and may provide increased hope for patients in this population.

Safety evidence

Blinatumomab has previously been approved for use in adults with relapsed or refractory Ph− BCP-ALL. As such, the safety profile of blinatumomab is well documented and the current approved indication is contingent on a risk management plan to inform providers and patients (and their caregivers) of the serious risk of cytokine release syndrome, neurological toxicities, and medication errors. The AEs that occurred during treatment in the MRD+ BCP-ALL population were consistent with the known safety profile of blinatumomab and no new safety signals were observed.

Blinatumomab has a favourable benefit–risk ratio for the treatment of patients with MRD+ BCP-ALL in haematological CR who would otherwise be at increased risk of haematological relapse, death, and HSCT failure. Patients with haematological relapse face an imminent risk of death, HRQoL decrement, and substantial costs. As described above, no other treatment options exist that are specifically indicated for MRD+ BCP-ALL patients in haematological CR; blinatumomab is therefore expected to become the SoC for this population.

Strengths of the clinical evidence base

A high rate of complete MRD response was observed in a trial designed with MRD-based inclusion criteria and MRD response as the primary endpoint; furthermore, a high rate of complete MRD response was observed across all MRD+ BCP-ALL subgroups, including patients in first CR. A clear difference in survival outcomes between MRD responders and non-responders was observed, supporting the value of achieving MRD negativity.

No new safety signals were observed and the comparative analysis with the MRD+ historical cohort demonstrated that blinatumomab treatment is associated with RFS and OS benefits compared with SoC chemotherapies: a complete MRD response achieved with blinatumomab is associated with an improvement in RFS and OS.

No other treatment options are specifically indicated for patients with MRD+ BCP-ALL, and the clinical evidence base presented establishes blinatumomab as a paradigm shift in therapy which is expected by clinical experts to become the new standard of care for this rare and deadly disease. The similarities of the results presented for the analyses censored at HSCT or post-blinatumomab chemotherapy and the uncensored analyses support the achievement of MRD negativity as an independent predictive factor for improved outcomes. Although single-arm trials, which present acknowledged limitations for the evidence base, both BLAST and MT103-202 were assessed to be of high quality, using appropriate methods for data collection (Section B.2.5) and the outcomes assessed in the study represented standard and objective outcome measures for the assessment of ALL therapies.

Weaknesses of the clinical evidence base

The key limitation of the evidence base is the lack of randomised controlled trials to inform relative efficacy estimates with blinatumomab. The single-arm nature of the BLAST and MT103-202 trials means that any ‘placebo effect’ resulting from the receipt of an active intervention (irrespective of the biological activity of that agent) cannot be adequately accounted for, reducing reliability of study results as a true estimation of treatment effect. Single-arm studies are more susceptible to selection and assessment bias, which may further reduce confidence in study results.

However, whilst RCTs represent the current 'gold standard' of trial design, it is not always possible or appropriate to conduct such a trial and single-arm studies may in some cases be the most appropriate form of study design. In cases where there is no clear standard of care or where currently available therapies have limited evidence of efficacy, it would be inappropriate and unethical to randomise patients to placebo and a single-arm trial is the appropriate choice of trial design. Given the high response rate observed in MT103-202, the well-established natural history of the disease and the low numbers of potentially eligible patients, the regulator agreed that a single-arm trial design was appropriate. A single-arm study design was chosen for BLAST on the basis that there was no effective standard available therapy for patients with MRD+ BCP-ALL.

Because of the single arm trial design an historical comparator study was required, which included a relatively small number of patients all of whom were in CR1 in contrast to the mix of patients recruited to BLAST (which covers the full anticipated license). Limitations which could have affected the comparability of the historical cohort to the BLAST population included differences in the MRD testing methods and timings between study sites, differences in local treatment protocols, the effect of HSCT status on treatment outcomes, and prior treatments received. However, the study was designed specifically to provide a population of patients closely matched to the BLAST study, and ultimately provided a population of patients that was highly generalisable to both BLAST and UK clinical practice. A comparison between BLAST and the historical comparator study was performed using propensity score analysis for patients in CR1. Both ATT and ATE weighting methods were used in order to consider the effect of any difference in these populations; both methods demonstrated the efficacy of blinatumomab, but the ATT weights were considered to better represent the population of the anticipated license.

B.2.13.2 Relevance of the clinical evidence base to the decision problem

Patient population

The patient population included in the BLAST study represents the anticipated licensed indication under consideration in this appraisal and provides the most relevant evidence for its use. The historical comparator study provides an appropriate historical control as to the effectiveness of the previous standard of care in the BLAST study locations and is sufficient to allow for comparisons clearly demonstrating efficacy to be made between BLAST and the previous standard of care. It may be noted, however, that the historical comparator trial, as well as recruiting a somewhat younger on average population, recruited a slightly more restricted population (those at CR1 only) than the BLAST trial. It may be considered that ATT weights (focussed on the BLAST population) rather than ATE weights (weighting between BLAST and historical comparator populations) are more appropriate when interpreting the results of the propensity score analysis.

Intervention

The BLAST trial provides evidence for the use of blinatumomab in the anticipated licensed indication and was conducted in European and UK centres which are considered to provide evidence generalisable to UK NHS treatment.

Comparators

The historical comparator study was designed to provide evidence for the standard of care in the centres in which the BLAST study was conducted. The protocol aimed to provide a cohort of

patients who were closely comparable to those recruited in the BLAST study to allow for propensity matching analyses to provide estimates of relative efficacy for blinatumomab. The approach of a single-arm study compared to an historical cohort was agreed with the regulator as appropriate for this rare disease for which no specific treatments have previously been available. The key difference of note between the historical cohort and the anticipated licensed indication is that the historical cohort included only patients in CR1, who are expected to have a more favourable outcome on the previous standard of care than those in CR2 and beyond.

Outcomes

The key outcomes measured in the trials, namely cytological complete response or relapse, molecular complete response or MRD, and mortality are all objective measures of disease in ALL and are unlikely to be affected by biases introduced by the study designs. As the experience of adverse events and quality of life both involve more subjective judgements it is possible that the open label, single-arm trial design may influence the observed outcomes, however it is not expected that this will introduce significant uncertainty into the overall interpretation of the evidence, given the clear and objective demonstration of efficacy observed in the trial.

B.2.13.3 End-of-life criteria

The evidence presented in Sections B.2.6 and B.2.9 demonstrates that blinatumomab for the treatment of MRD+ BCP-ALL patients meets the criteria for a life-extending end-of-life treatment; as described in Table 4, median OS in the historical comparator study was [redacted], while for patients treated with blinatumomab, median OS was [redacted] after more than 40 months follow up (Section B.2.9.4). Compared to SoC chemotherapy, blinatumomab was found to reduce the risk of death at 18 months by [redacted]. It should be noted that in the historical control patients treated with SoC chemotherapy, a small number of patients were observed to survive for a long time. Given the skew caused by this small group of patients, it was considered appropriate to use median OS values, rather than the mean, so as to more accurately represent the patient population as a whole. This skew effect and use of median OS rather than the mean has been noted in previous appraisals where the Committee agreed that consideration of medians was more appropriate.[74]

Furthermore, given that blinatumomab is indicated for a rare condition in a very small number of patients (85 per year) who have a huge unmet medical need and who stand to gain substantially from access to blinatumomab, this therapy meets many of the criteria for appraisal under the HST framework. Consequently, blinatumomab should be evaluated taking into account a wider range of criteria about the benefits and costs.

Table 32: End-of-life criteria

Criterion	Data available	Reference in submission (section and page number)
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	Median OS, using ATT-weighted propensity score matching analyses (Section B.2.9.4) for SoC chemotherapy was [redacted]. The estimated mean survival (undiscounted) in the economic analysis was almost 5x	B.2.9.4, pages 74–75

	greater than the median survival ([redacted] years) in the SoC arm; however, this is reflective of the small proportion of patients who achieve long-term survival (~20%). For this reason, the median survival is considered to be a more suitable representation of the anticipated survival in the patient population as a whole.	
There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment	Median OS, using ATT-weighted propensity score matching analyses (Section B.2.9.4), was not estimable after more than 40 months follow-up for blinatumomab thus demonstrating a >20 month OS survival benefit when compared to SoC. The estimated mean survival (undiscounted) in the economic analysis was [redacted] years in the blinatumomab arm, resulting in an incremental survival benefit of [redacted] years	B.2.9.4, pages 74–75

Abbreviations: OS: overall survival; ATT: average treatment effect on the treated; SoC: standard of care; CI: confidence interval; N.E.: not estimable; NHS: National Health Service.

B.2.13.4 Conclusion

- Blinatumomab is the first and only drug indicated specifically for MRD+ BCP-ALL patients in haematological CR (Section B.1.3).
- Blinatumomab achieves MRD negativity in 78% of MRD+ BCP-ALL patients in haematological CR within the first cycle (Section B.2.6.1).
- Compared with SoC, blinatumomab reduces the risk of haematological relapse or death by 58% and more than quintuples the median RFS in MRD+ BCP-ALL patients in haematological CR (Section B.2.9.4).
- Compared with SoC, blinatumomab reduces the risk of death by [redacted] and [redacted] the median OS in MRD+ BCP-ALL patients in haematological CR (Section B.2.9.4).
- No new safety signals were observed in patients pooled from the MRD studies, beyond the existing safety profile of blinatumomab (Section B.2.10).

B.3 Cost effectiveness

B.3.1 *Published cost-effectiveness studies*

No published cost-effectiveness studies were identified for blinatumomab in patients with MRD+ BCP-ALL in haematological CR. For details of the economic SLR, please see Appendix G.

B.3.2 *Economic analysis*

A partitioned survival analysis (PartSA) model was used to estimate expected RFS, OS, lifetime costs of ALL treatment, and quality-adjusted life years (QALYs)s in patients with Ph- MRD+ B-precursor ALL in haematological CR. The model was developed in Microsoft Excel. RFS, OS, duration of treatment with blinatumomab, probabilities of HSCT, and utility values were based on data from the BLAST and historical control studies and other sources. Because this is the first economic evaluation in this disease area, a *de novo* economic model was required. The modelling approach is similar to that used in the manufacturers submission to NICE in response to the STA of blinatumomab in relapsed/refractory (R/R) Ph- B-precursor ALL.[75]

When possible, data from BLAST (MT103-203 Study) and the historical control (20120148 Study) were based on the 73 patients from BLAST and the 182 patients from the historical control study who were included in the PAS of the propensity-matched comparison of patients in the MT103-203 and 20120148 studies. This population includes patients in BLAST and the historical comparator study meeting the following criteria:

- Ph- B–precursor ALL;
- First complete haematological remission (CR1);
- MRD+ at a level of >1 x 10^{-3};
- ≥18 years old at MRD positivity (20120148 Study) or first blinatumomab treatment (MT103-203);
- Complete baseline covariate set;
- Time to relapse greater than 14 days from MRD detection (20120148 Study).

When possible, data for patients in the historical control were weighted using average treatment effect among treated patients (ATT) weights calculated based on the inverse probability of treatment weights (IPTWs) from the propensity matched comparison of BLAST versus the historical control. ATT rather than average treatment effect (ATE) weights were used as results based on ATT weight can be generalised to the population of patients in BLAST rather than the combined populations of the BLAST and historical control studies. This is appropriate as the historical control study was designed *a priori* to match patients in BLAST, and the use of propensity score analysis was conducted only to control for residual confounding after matching. Although ATT weights were used for base-case analyses, a scenario analysis was conducted using ATE weights.

The cost of blinatumomab was based on the discounted price offered to the NHS through its approved PAS. Other costs were based on NHS reference costs and published studies. The incremental cost-effectiveness ratio (ICER) was expressed in terms of the incremental cost per QALY gained.

B.3.2.1 Patient population

The cost-effectiveness of blinatumomab was evaluated in patients with Ph- B-precursor ALL in CR1 with MRD positivity. This population represents a subgroup of patients in the BLAST trial, and is narrower than the anticipated license considered in the submission. This population is appropriate for the economic evaluation as it is expected that blinatumomab will be used as early as possible in the treatment pathway where the benefits of treatment are likely to be greatest. Given the lack of data for a historical control in patients in second or subsequent haematological relapse (CR2), evaluation of cost-effectiveness in the broader set of patients was infeasible. It should be noted that in the NICE guidance TA450 recommending blinatumomab for previously treated Ph- ALL, the appraisal committee concluded that, among patients with R/R ALL, those with no salvage therapy would be the most relevant population for the appraisal as it is consistent with where the product will most likely be used in routine clinical practice.[75] The general approach of focusing on early use in the treatment pathway is therefore consistent with the appraisal committee conclusions in the prior assessment of blinatumomab in R/R ALL.

B.3.2.2 Model structure

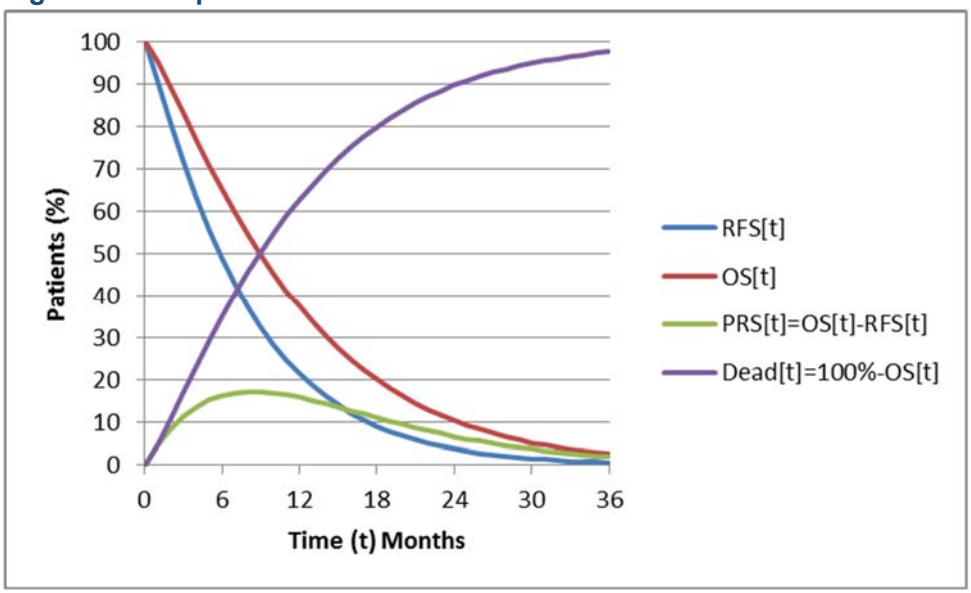
The model is implemented as a Microsoft Excel workbook and uses a PartSA approach with states defined based on relapse and death. PartSA is a transparent, intuitive approach which yields estimates of survival that correspond closely to survival observed during the study that are the basis for the evaluation.[76] The PartSA approach has been used in numerous prior economic analyses of treatments for oncology therapies including haematologic malignancies,[76] and in the recent manufacturer's submission in response to the STA of blinatumomab in R/R B-precursor ALL.[75]

Despite its strengths, limitations of this approach should be noted. Unlike state transition models (e.g., Markov cohort models or patient-level simulations), the PartSA approach does not permit explicit modelling of dependencies among clinical events such as MRD response, relapse, allogeneic HSCT, receipt of salvage therapy and survival. While it is theoretically possible to build a state transition model that could incorporate the relationships between these various endpoints, the estimation of such models in this instance would face a host of challenges including limited sample size available for modelling due to the rarity of the disease of interest, potentially incorrect specification of the distributions of conditional (i.e., transition) probabilities, and the likely need for additional assumptions regarding the nature of dependencies and the effects of treatment, as well as potential biases in the estimation of transition probabilities due to informative censoring.[77] Given the relative importance of OS in determining the economic value of blinatumomab in this indication, and the availability of survival for patients receiving SoC from the historical control study out to approximately eight years, the strength of the partitioned survival model in fitting to observed survival data outweighs the potential theoretical benefits of a structural model of OS based on interim endpoints. To address concerns regarding biases associated with extrapolation of survival projections, model projections of RFS and OS were compared against external data on long-term survival by MRD response. This approach is consistent with recent recommendations for the conduct of economic evaluations based on PartSA by the NICE Decision Support Unit (DSU).[76]

The PartSA model used in the economic analysis includes three states: relapse-free (RF), post relapse (PR) and dead. All patients are assumed to enter the model in the RF state. During the course of the modelling time horizon, patients may experience relapse and enter the PR state or die and enter the death state. Relapse is defined as haematological relapse as described in the

BLAST trial. Patients who are alive are divided (or "partitioned") according to relapse status under the assumption that relapse has implications for HRQoL and costs. Membership in the three states over time is determined by survival curves for RFS and OS. RFS provides the proportion of patients remaining in the RF health state over time. Membership in the dead state is calculated as the complement of the OS curve (i.e., one minus OS) at each point in time. Membership in the PR state is calculated as the difference between OS and RFS at each time point. The process of deriving membership in the RF state and the dead state (PR[t]) and Dead[t], respectively) is illustrated in Figure 19.

Figure 19. Simplified schematic of PartSA model



Abbreviations: RFS: relapse-free survival; OS: overall survival; PRS: post-relapse survival.

HRQoL is assumed to be conditional on health states and, for patients in the RFS state, on versus off blinatumomab treatment, and MRD response. Costs of follow-up and monitoring also are assumed to depend on health states. Costs of blinatumomab therapy are modelled independently of health states. Patients receiving SoC maintenance therapy are assumed to receive it for a maximum of 2 years or until relapse, allogeneic HSCT, or death, whichever occurs first. Because data on the incidence of allogeneic HSCT in BLAST and the historical control study were reported only up to relapse, costs of HSCT occurring before and after relapse are considered separately. Those occurring before relapse are modelled independently of relapse (i.e. the probabilities of receiving pre-relapse HSCT are not contingent on RFS).

For the purpose of calculating the costs of HSCT and salvage therapy, the model accounts for the proportion of RFS events that are deaths versus relapses. Those for whom the RFS event is relapse are assumed to (potentially) incur the costs of post-relapse HSCT and salvage treatment. For the purpose of discounting, the costs of salvage therapy and post-relapse HSCT is assumed (for convenience) to be incurred at relapse (this assumption is reasonable because most patients who receive salvage will get it soon after relapse and those who receive HSCT will likely get it soon after response to salvage treatment). Those who die are assumed to incur ALL-related terminal care costs.

It should be noted that in the base-case, it was assumed that salvage therapy for patients experiencing relapse would be SoC chemotherapy, consistent with the salvage therapy received

among patients in the BLAST trial and the historical control study. While the model estimates of OS and RFS are internally consistent with this assumption, it does not reflect the likely use of blinatumomab in patients with R/R ALL as salvage therapy given the recent NICE guidance.[75] Because modelling of a "counterfactual" scenario in which relapsing patients would receive blinatumomab was not feasible using the data from BLAST and the historical control study within the partitioned survival model structure, a scenario analysis was conducted in which incremental costs and QALYs generated by the model were adjusted to reflect the difference between treatments in the percent of patients receiving blinatumomab as salvage therapy and the incremental costs and QALYs associated with blinatumomab versus SoC chemotherapy salvage for patients without prior salvage therapy based on the economic model used in the evaluation of blinatumomab versus SoC salvage therapy in R/R ALL.[75] This scenario is described in greater detail in Section B.3.8 below.

In the base-case as it was assumed that the parametric survival distributions used in the model would accurately reflect the short and long-term impact of blinatumomab on RFS and OS. Scenario analyses were conducted in which it was assumed that after some defined period of time ("duration of benefit"), patients receiving blinatumomab will have the same RFS and OS hazards as patients receiving SoC therapy.

To account for the long-term age-related increase in non-ALL related mortality, estimated survival distributions for RFS and OS are combined with age- and sex-matched general population mortality estimates. In particular, the model uses the maximum of the probability of death from the RFS and OS distributions and the general population mortality, adjusted to reflect the potential long-term effects of complications of cytotoxic chemotherapy, and/or allogeneic HSCT on survival. Thus, the adjusted general population mortality is used as a floor below which the mortality in the model may not fall. To calculate age- and sex-matched mortality, mean age at therapy initiation was assumed to be 45.4 years and 56.2% of patients were assumed to be male, based on the demographic characteristics of the CR1 population in BLAST.[56]

In the base-case, it was assumed that patients who remain alive after five years no longer incur ALL-related costs (other than follow-up care for allogeneic HSCT received previously) and have HRQoL consistent with age- and sex-matched general population norms, adjusted to reflect the potential long-term effects of complications of radiotherapy, cytotoxic chemotherapy, and/or allogeneic HSCT on HRQoL. This assumption is based on clinical expert opinion that patients who survive for five years are likely to be cured of ALL, but may have residual decrements in HRQoL due to prior treatments.

For HRQoL, data on the decrement in utility values associated with long-term effects of complications of radiotherapy and cytotoxic chemotherapy are unavailable. Accordingly, it was assumed that utility values in long-term ALL survivors would, at best, rebound to mid-way between the utility value for patients with MRD response following blinatumomab therapy and the age- and sex-matched general population utility value. Accordingly, when applying the general population utility values, these values were adjusted by a constant absolute decrement equal to one half the difference in the age- and sex-matched general population utility value and the estimated utility value for patients with MRD response after blinatumomab therapy.

For mortality, it was assumed in the base-case that the probability of death is never less than 4-fold greater than the age- and sex matched general population mortality. This assumption is based on several considerations. Data on long-term (i.e., >10 years) excess mortality for adult patients with MRD+ Ph- B-precursor ALL in first haematological CR receiving blinatumomab or

SoC maintenance therapy in the modern era are unavailable. Because many long-term survivors are likely to have received allogeneic HSCT, data on long-term survivorship for transplant survivors may be an appropriate proxy for estimating long-term survivors in the model. Numerous studies have examined the long-term excess mortality of patients undergoing stem cell transplants and have reported that although mortality rates decline dramatically during the first few years after transplant, they remain elevated for many years after transplant.[78-83] In the NICE appraisal consultation for inotuzumab ozogamicin, the appraisal committee's preferred assumption was a 4-fold increase in mortality versus general population 3 years after stem cell transplant.[84, 85] This assumption was based on data from study by Martin et al. (2010) of 2,574 patients who survived without recurrence of the original disease for at least 5 years after allogeneic or autologous haematopoietic cell transplantation from 1970 through 2002 at Fred Hutchinson Cancer Research Center and the University of Washington, Seattle, WA.[78] These authors reported a four- to nine-fold increase in mortality relative to the general population after 15 years. The use of the value from the lower end of this range was based on the acknowledgement of the evidence review group that the data on long-term survival from this and other studies are based on cohorts of patients who received transplant decades ago and therefore may overstate long-term mortality for patients receiving transplant in current practice. In the case of the Martin study, the mortality ratio versus the general population declined to 4.0 until approximately 15 years after which it increased to 9.0 by 25-30 years. However, the person time during this period is heavily weighted towards patients who received transplant early in the study and for whom outcomes are likely to be less favourable. Hence, the significance of the increased risk in years 15–30 reported by Martin is highly uncertain. Additionally, patients receiving transplant in earlier stages may experience better outcomes than those receiving it in later stages. Wingard and colleagues (2011) reported that stage at transplant was a predictor of worse survival in ALL patients who survived ≥ 2 years after allogeneic HSCT (multivariate relative risk = 1.77 for late versus early).[79] Socié et al. (1999) reported that that among patients with ALL who survived two years after transplant, the HR of death for patients in CR2 was 1.75 versus those in CR1.[82] Based on these considerations, a 4-fold increase in mortality was assumed, consistent with that assumed for post-transplant patients in the NICE appraisal consultation for inotuzumab ozogamicin.

Key features of the economic analysis are summarised below in Table 33.

Table 33: Features of the economic analysis

	Current appraisal	
Factor	**Chosen values**	**Justification**
Modelling approach	PartSA	PartSA is a transparent, intuitive approach which yields estimates of survival that closely correspond to those observed during the trial. The PartSA approach has been used in prior economic analyses of treatments for haematological malignancies including that in the recent manufacturer's submission for blinatumomab in R/R B-precursor ALL. Data on survival for patients receiving SoC from the historical control study was available out to approximately 8 years, limiting the

		need to develop a structural model of OS based on interim events.
Time horizon	50 years	A 50-year time horizon corresponds to a lifetime projection for a typical patient in the CR1 population of BLAST. The mean age of blinatumomab patients in the propensity matched analysis of BLAST and the historical control was 45.4 years. After 50 years, under base-case assumptions, approximately 98% of patients in the SoC group are projected to be dead.
Cycle length	1 week	A weekly cycle length was used to permit accurate estimation of survival without the need for half-cycle correction.
Treatment waning effect?	None	It is assumed that the parametric survival distributions fit to the data on RFS and OS for the matched populations of BLAST and the historical control study capture any waning of the treatment effects of blinatumomab on these endpoints. The use of limited duration treatment effects was examined in scenario analyses.
Source of utilities	Utility values for the RFS state were based on data on EQ-5D utility value from the BLAST trial using UK tariffs. Utility values for patients in PR were based on EQ-5D utility values mapped from the EORTC QLQ-C30 among for patients receiving SoC chemotherapy in the TOWER trial. Utility values for long-term survivors were based on general population norm utility values, adjusted for a long-term decrement in utility due to exposure to cytotoxic chemotherapy and HSCT.	Use of EQ-5D utility values is consistent with the NICE reference case.[86] Data on HRQoL for patients receiving SoC maintenance were unavailable and were based on estimates of EQ-5D utility values among patients in BLAST who were off therapy by MRD response. Because post-relapse utility assessments in BLAST were limited in number, post-relapse utility values were based on utility values for patients receiving SoC salvage therapy in the TOWER trial of blinatumomab as salvage therapy for R/R B-precursor ALL.
Source of costs	Estimates of exposure to blinatumomab and receipt of HSCT were from BLAST and the historical control study. The price of blinatumomab was based on the discounted price offered to the NHS through an approved PAS. Cost of HSCT was from published sources. State dependent healthcare resource use was based on face-to-face interviews of UK clinicians. Costs of other medications were from the BNF or	Consistent with the NICE reference case.[86]

	eMIT. Other unit costs were based on NHS reference costs.	
Discounting	3.5% for costs and effective (1.5% for health outcomes in sensitivity analysis).	Consistent with the NICE reference case.[86]

Abbreviations: R/R: relapsed or refractory, ALL: acute lymphoblastic leukaemia; SoC: standard of care; OS: overall survival; CR1: first complete response; RFS: relapse-free survival; EQ-5D: EuroQol Five Dimensions; EORTC-QLQ-C30: quality of life of cancer patients questionnaire; UK: United Kingdom; BCP: B-cell precursor; HSCT: haematopoietic stem cell transplantation; MRD: minimal residual disease; NHS: National Health Service; PAS: patient access scheme; NICE: National Institute for Health and Care Excellence; BNF: British National Formulary; eMIT: Drugs and pharmaceutical electronic market information.

B.3.2.3 Intervention technology and comparators

The intervention of interest is ≤ 4 consecutive cycles of blinatumomab 28 µg/day at a constant IV infusion over 28 days followed by an infusion-free interval of 14 days, with treatment stopped following haematological relapse and with patients suitable for HSCT after ≥1 cycle of blinatumomab possibly receiving allogeneic HSCT instead of further cycles of blinatumomab. This treatment is consistent with that employed in BLAST with the exception that blinatumomab will be assumed to be administered at 28 µg/day, in line with the currently approved dosage, rather than the 15 µg/m²/day dosage employed in BLAST. Consistent with protocol of the BLAST trial, patients receiving blinatumomab were assumed to receive IT triple combination CSF prophylaxis consisting of 4 mg of dexamethasone, 15 mg of methotrexate, and 40 mg of cytosine arabinoside (cytarabine) every 3 months until relapse or receipt of HSCT for two years. Although the protocol allowed for continuing CSF prophylaxis after 2 years based on investigator discretion, it was assumed to be limited to two years in order to be consistent with that assumed for maintenance therapy (see Section B.3.5).

The only comparator of interest is SoC treatment, which is assumed to be conventional maintenance chemotherapy for MRD+ patients in haematological CR. Based on the maintenance regimen for non-transplant patients used in the UKALL14 trial, a randomised phase III trial of SoC chemotherapy with or without rituximab, and with or without nelarabine, in patients with newly-diagnosed ALL,[45] SoC maintenance therapy was assumed to include the following:

- Vincristine 1.4 mg/m^2 (max 2 mg/dose) IV every 3 months for 2 years
- Prednisolone 60 mg/m^2 orally 5 days every 3 months for 2 years
- Mercaptopurine 75 mg/m^2 orally daily for 2 years
- Methotrexate 20 mg/m^2 orally once per week for 2 years
- CSF prophylaxis with intrathecal methotrexate 12.5mg every 3 months for 2 years

Maintenance therapy was assumed to be discontinued upon relapse or receipt of SCT.

In BLAST, CR was defined as having < 5% blasts in bone marrow after at least 3 intense chemotherapy blocks. Accordingly, it is appropriate to assume that patients receiving blinatumomab will have already received induction, intensification, and consolidation therapy, and that maintenance therapy is therefore the appropriate comparator. This approach may be conservative, however, as it is anticipated that blinatumomab will be used in the post-induction setting, and therefore may replace the use of intensification and consolidation therapy as well as maintenance treatment.

B.3.3 *Clinical parameters and variables*

B.3.3.1 Relapse-free and overall survival

Because survival distributions for RFS and OS in BLAST were incomplete, it was necessary to extrapolate survival distributions beyond the end of the trial to obtain unbiased estimates of the gains in life expectancy and QALYs with blinatumomab. This extrapolation was performed by fitting parametric models to individual patient data on RFS and OS from the patients in the propensity-matched analysis of patients from BLAST and the historical comparator study based on the ATT-weights.

The fitting of parametric models was performed using Flexsurv, an R package for fully-parametric modelling of survival.[87] A wide range of parametric distributions were considered, including the exponential, Weibull, log-logistic, lognormal, Gompertz, gamma, and restricted cubic spline (RCS) distributions.

For each time-to-event outcome (i.e. RFS and OS) and distribution, models were estimated alternately (a) including a single indicator variable for treatment group in the model formulation ("restricted models") and (b) including treatment-group interaction terms for every distributional parameter ("unrestricted models"), as shown in Table 34. With both approaches, the distributions of survival for the treatment and control group are assumed to be of the same class (e.g., both are Weibull). However, with the first approach (restricted models), the effect of treatment is restricted to a single distributional parameter (e.g. the scale parameter of the Weibull distribution) and yields projections of survival that are consistent with proportional hazards, accelerated failure time, or other univariate treatment effect models, depending on the distribution (e.g. the Gompertz is a proportional hazards model, the lognormal and log-logistic are accelerated failure time models, and the exponential and Weibull are both proportional hazards and accelerated failure time models). The second approach (unrestricted models) places no such restrictions on the distributional parameters or the assumed nature of treatment effect within the class. Estimating these restricted and unrestricted models in this way permits comparison of the Bayesian information criterion (BIC, and other fit statistics) for unrestricted and restricted models (which would not be possible if the unrestricted models were estimated as two separate regression equations – one for each arm of the trial). The assumption that the distributions of survival for the treatment and control group are of the same class is reasonable because any differences in shapes between arms can generally be accommodated by the use of unrestricted forms of more flexible survival distributions (e.g. RCSs).

Table 34. Alternative parameterisations of the treatment effect employed in parametric non-cure models

Treatment Effect	Restricted	Unrestricted
HR, acceleration factor (AF), or other treatment effect parameter	✓	✓
Other parameters	×	✓

Abbreviations: HR: hazard ratio; AF: acceleration factor.

In addition to the parametric models described above, parametric cure models were fitted using the stsrsmix and strsnmix Stata procedures,[88] in order to account for the potential that a

significant subset of patients who might experience long-term RFS or OS. The use of parametric cure models is appropriate under such circumstances.[89] Both mixture and non-mixture cure models were considered. Mixture cure models treat the patient population as a mix of cured and uncured patients, where cured patients are at no excess risk for the event and uncured patients face excess risk as modelled by a simple parametric survival distribution (e.g. Weibull, lognormal, gamma). By contrast, non-mixture models rescale a simple parametric survival model such that survival asymptotically approaches the estimated cure fraction. Both mixture and non-mixture models were run using Weibull, lognormal, and gamma as baseline distributions.

For each time-to-event outcome (i.e. RFS and OS) and distribution, parametric cure models were estimated alternately (a) including a single indicator variable for treatment group which varied the cure fraction in the model formulation ("cure"), (b) including treatment group interaction terms which varied the cure fraction and a single parameter of the baseline distribution ("cure + restricted"), and (c) including treatment-group interaction terms for every distributional parameter ("cure + unrestricted"), as show in Table 35.

Table 35. Alternative parameterisations of the treatment effect employed in parametric cure models

Treatment Effect	Cure	Cure + Restricted	Cure + Unrestricted
Cure probability	✓	✓	✓
HR or AF	×	✓	✓
Other parameters	×	×	✓

Abbreviations: HR: hazard ratio; AF: acceleration factor

Selection of parametric distributions for RFS and OS was based on several factors, including internal consistency, statistical fit, visual fit, evidence related to the underlying treatment effect model, and consistency with available external data. In order to leverage the larger number of events observed and ensure internal consistency, the base-case RFS distribution was selected first and then used to inform the selection of an OS distribution.

Internal consistency between RFS and OS distributions was assessed in two ways. First, OS distributions were considered inconsistent with a given RFS distribution if they crossed at any time during the model projection, since this would present a logical inconsistency. While it would be possible to resolve this consistency by simply setting RFS to the minimum of selected RFS and OS distributions, this would imply that, from that time forward, no patients would be remaining alive following their first relapse. Such a result would be inconsistent with clinical opinion previously accepted by the appraisal committee for the NICE STA of blinatumomab in Ph- R/R ALL, which accepted the proposition that some R/R patients are effectively cured.[75] If a subset of relapsing patients achieve long-term survival, then RFS should remain below OS throughout the model projection. Based on these factors, OS distributions which crossed the selected RFS distribution were not considered. Second, OS distributions were preferred if the difference in expected post-relapse survival (PRS) for blinatumomab and SoC was relatively small, as the benefits of blinatumomab in patients in CR1 and MRD+ on PRS are more uncertain.

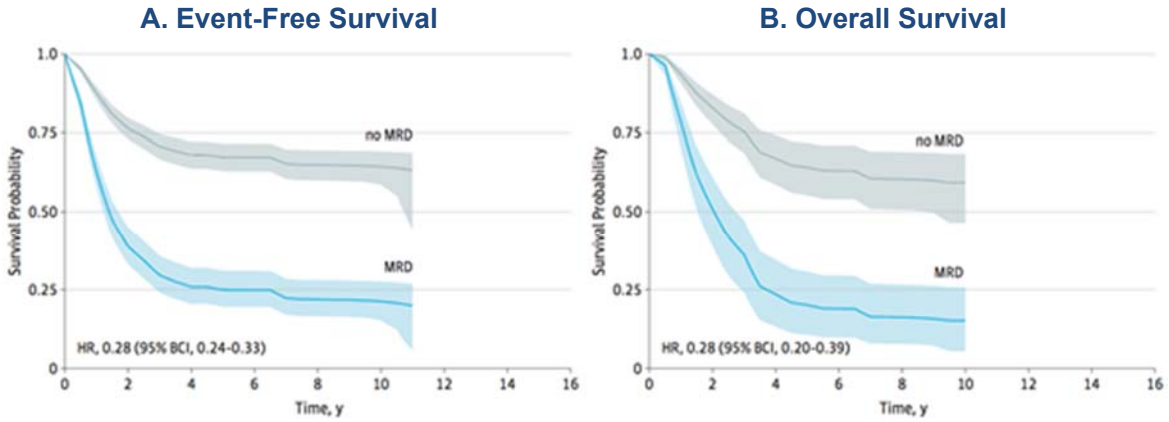
With respect to statistical fit, the BIC statistic was used as the primary fit statistic since it penalises overly complex models and its use mitigates the risk of overfitting statistical noise in the tails of the observed distributions.

Diagnostic plots for the nature of treatment effects were produced using an extension of an approach proposed by Bagust and Beale and in accordance with recommendations from the NICE DSU Technical Support Documents on survival analysis.[90, 91] With this extended approach, an estimated treatment effect for each of four different treatment effect assumptions (i.e. constant shift in survival time, accelerated failure time, proportional hazards, and proportional odds) was applied to failure times in the control group to obtain a counterfactual Kaplan-Meier (K-M) survival distribution for the control group reflecting the expected outcome had those patients received study treatment with the specified treatment effect assumption. The counterfactual control group survival distribution was then compared with the observed survival distribution for the group receiving study treatment. If the treatment effect assumption is accurate, the two curves should overlap. This approach permits comparisons of different treatment effect assumptions on the same (natural) scale. Deviations from proportional hazards were also assessed based on the Schoenfeld residuals.

The external validity of RFS and OS distributions were based on comparisons of model projections with estimates of EFS and OS from Berry et al., the most recent and only meta-analysis of studies assessing the association between MRD status and clinical outcomes such as RFS and OS in adults with ALL (EFS in the study by Berry et al. was assumed to be approximately equivalent to RFS in BLAST and the historical control study).[34] Event-free survival (EFS) and OS by MRD response from this study are shown in Figure 20. Patients with MRD response had statistically significant improvements in EFS and OS, both in terms of the HR and the probability of survival at ten years, suggesting a large and durable effect of MRD response on outcomes. However, reported EFS and OS from Berry et al. (2017) were nearly identical, with EFS exceeding OS at ten years both for those with and without MRD response. This discrepancy may result from the different set of studies included in the analyses of EFS and OS by Berry et al. While this lack of internal consistency means that projections of EFS and OS based on Berry cannot be used without adjustment to inform model projections, they were still considered to contain information relating to the shape of RFS and OS curves, as well as the magnitude of benefit which might be expected based on MRD response.

Benchmark projections of EFS and OS were therefore obtained by weighting the MRD+ and MRD- patients in Berry based on estimated MRD response rates. For patients receiving blinatumomab, the proportion with MRD response was based on the percent of patients with MRD response among the 73 blinatumomab patients in the propensity-matched analysis of BLAST and the historical control study (83.6%). It should be noted that this percentage included 2 patients who achieved response in cycle 2 of BLAST (whereas the primary endpoint of BLAST only included patients with MRD response at the end of cycle 1). For patients receiving SoC maintenance therapy, the proportion of patients who might achieve a delayed MRD response is unknown, as this information was not examined in the historical control study, and has not been reported in the literature. Discussions with clinical experts indicate that this proportion is no greater than 10%. It was therefore assumed that 8% of patients receiving SoC maintenance therapy would achieve a delayed MRD response.

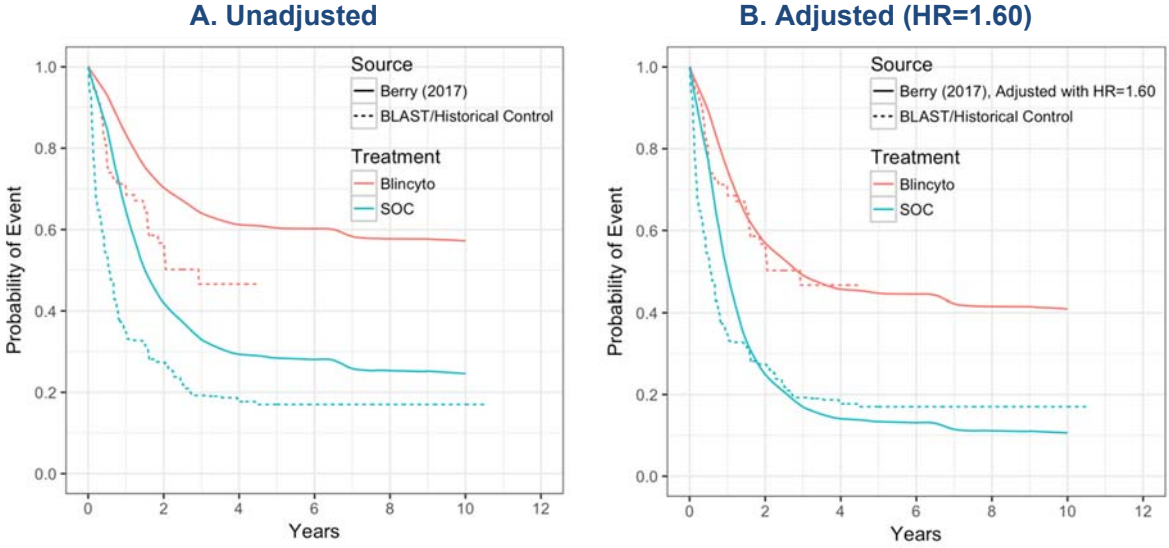
Figure 20. EFS and OS among MRD+ and MRD- ALL patients from meta-analysis by Berry et al.



Source: Berry et al. (2017)[34]

Projections of EFS based on Berry et al. are shown alongside RFS from BLAST and the historical comparator study in Figure 21a. Projections based on Berry consistently overestimated RFS compared to those from BLAST and the historical comparator study (potentially due to heterogeneity of population and definition of the end point as discussed above), but appeared to have a similar shape. In order to align projections of RFS for SoC with the historical comparator study for the purposes of visual analysis only, a HR of 1.60, chosen based on visual inspection was applied to the RFS estimates from Berry et al., as shown in Figure 21b.[34] The application of this same hazard also aligned projections for blinatumomab based on the study by Berry and BLAST, indicating the impact of MRD response on RFS as predicted based on Berry was aligned with what was observed in BLAST and the historical control study.

Figure 21. Projected RFS based on Berry et al. for blinatumomab and SoC versus BLAST and historical control study



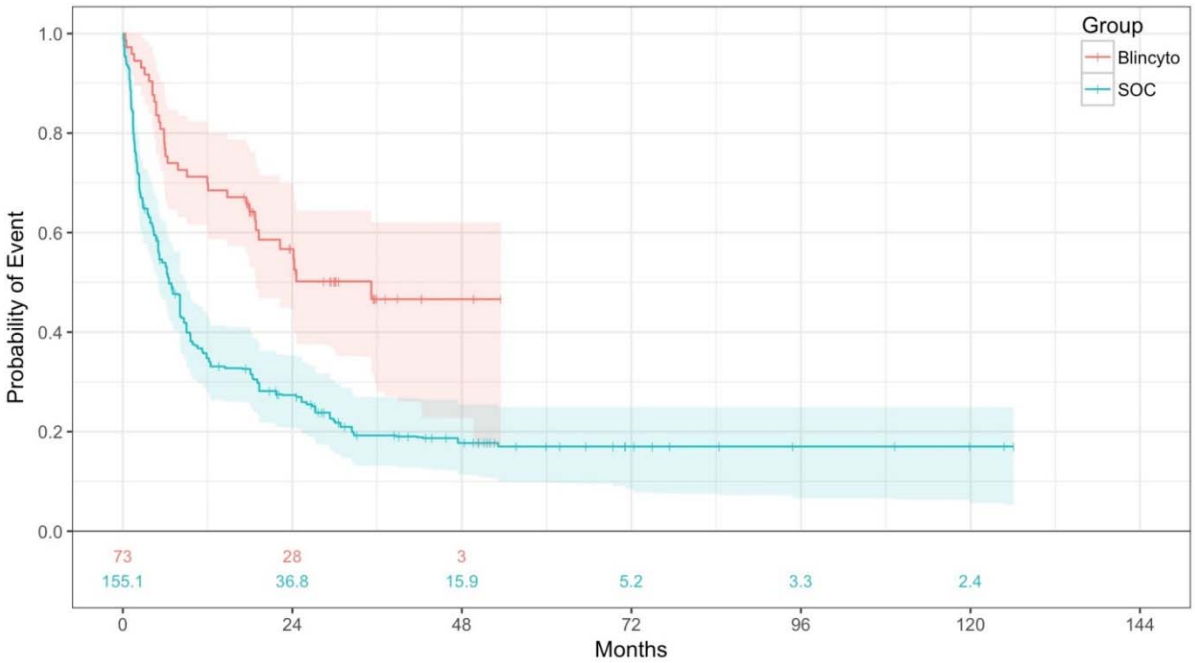
Footnotes: Adapted from Berry et al. (2017)[34]
Abbreviations: RFS: relapse-free survival; SoC: standard of care; HR: hazard ratio.

When considering OS, the overall shapes of the survival curves are also highly similar, and while projections yielded comparable estimates of OS for SoC to the historical control study, those for blinatumomab somewhat overestimated OS for blinatumomab relative to BLAST. This suggests, that the survival benefit of MRD response was larger in Berry than in BLAST and the historical comparator study, which could be due to the heterogeneity in various factors across the different studies contributing to the meta-analysis by Berry et al. including the patient populations, timing of MRD testing, and the threshold of MRD negativity, etc. Nevertheless, analysing the overall shapes of the curves provides a useful validation of survival for the population considered in the decision problem

Relapse-Free Survival

K-M estimates of ATT-weighted RFS in the CR1 subgroup are shown in Figure 22. Clear separation was observed between blinatumomab and SoC throughout the follow-up of BLAST. A general pattern of decreasing hazards was observed with long periods without events at the tail in both arms.

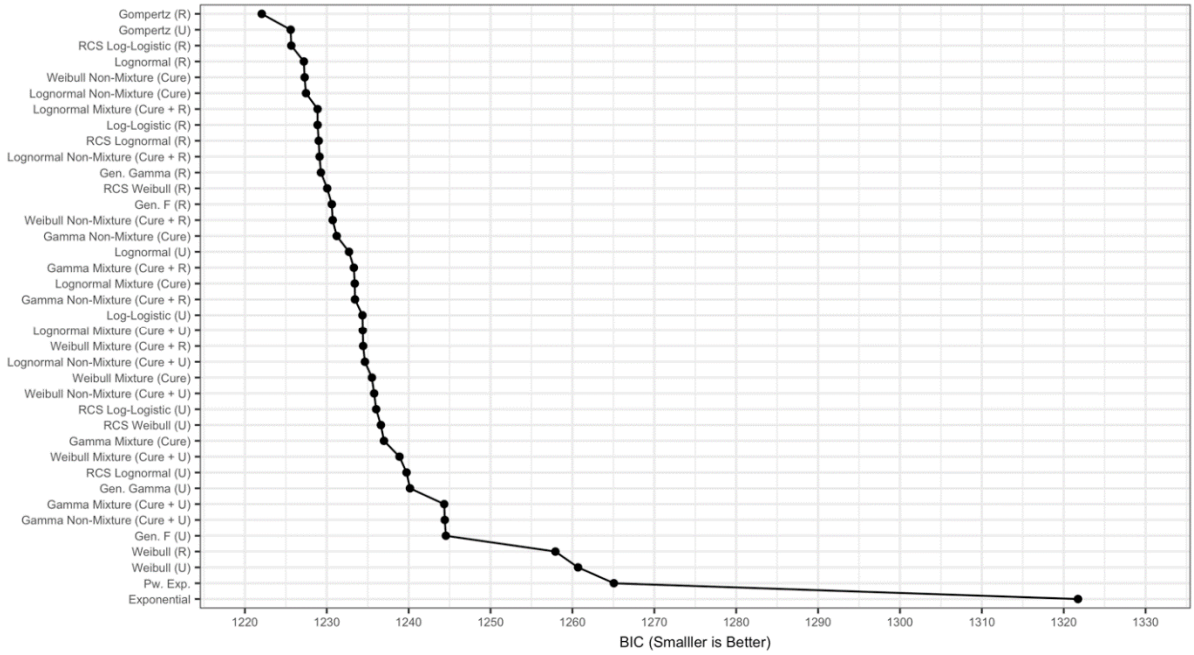
Figure 22. K-M estimates, ATT-weighted CR1 RFS



Abbreviations: ATT: average treated effect on the treated; CR1: first haematological complete response; RFS: relapse-free survival; SoC: standard of care.

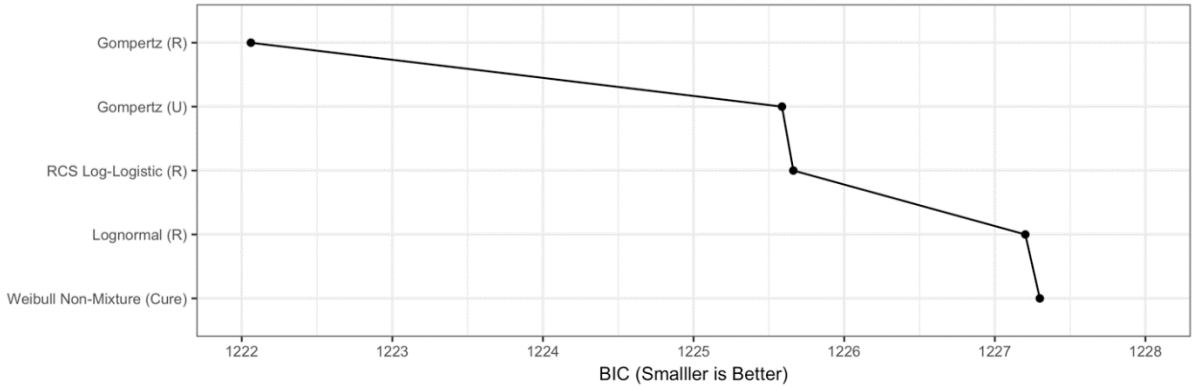
Statistical fit, as measured by BIC, is shown for all distributions in Figure 23. Statistical fit varied widely between distributions, with a score of 1,222 for the best fitting restricted Gompertz distribution and 1,321 for the worst-fitting exponential distribution. In order to focus on the best fitting distributions, only the top five distributions were considered. Additional information on all fitted RFS distributions can be found in Appendix P. Statistical fit for the top five distributions only is shown in Figure 24.

Figure 23. Fit statistics for parametric survival distributions fit to RFS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATT weights



Abbreviations: ATT: average treatment effect on the treated; CR1: first haematological complete response; RFS: relapse-free survival; R: restricted; U: unrestricted; RCS: restricted cubic spline; BIC: Bayesian information criterion.

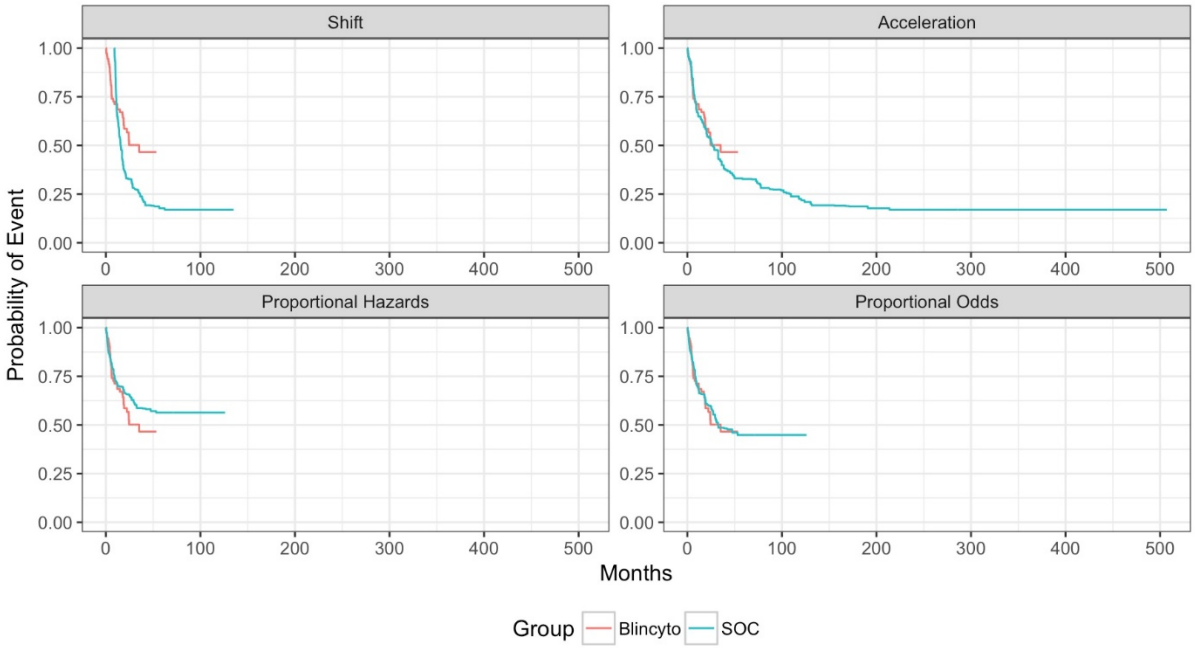
Figure 24. Fit statistics for five best-fitting parametric survival distributions fit to RFS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATT weights



Abbreviations: ATT: average treatment effect on the treated; CR1: first haematological complete response; RFS: relapse-free survival; R: restricted; U: unrestricted; RCS: restricted cubic spline; BIC: Bayesian information criterion.

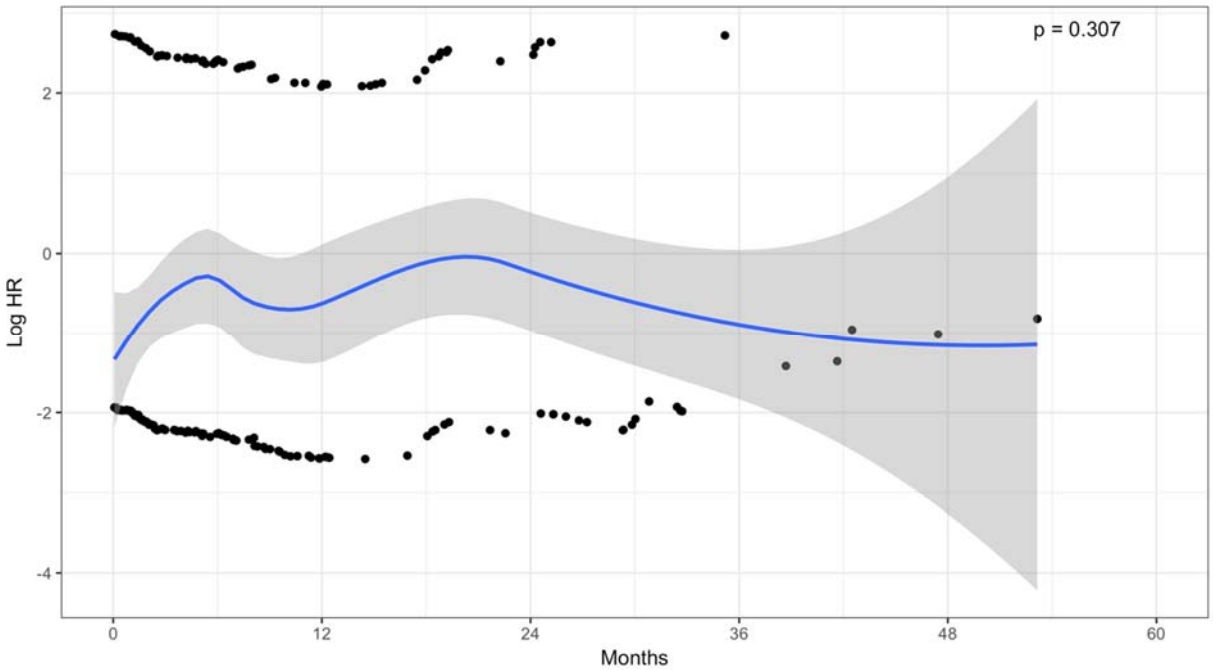
Treatment effect counterfactual plots are shown in Figure 25. RFS from BLAST and the historical comparator study were best represented by proportional odds and accelerated failure time, with no systematic bias observed. Proportional hazards performed well during the first 18 months of the study, but overestimated the magnitude of benefit during the remainder of the study. However, no significant deviation from proportional hazards was identified in the Schoenfeld residuals (Figure 26). A fixed shift in RFS was ruled out due systematic bias throughout the trial follow-up.

Figure 25. Treatment effect counterfactual plots for RFS for patients receiving blinatumomab and SoC from propensity matched analysis of BLAST and historical control study, ATT weights



Abbreviations: ATT: average treatment effect on the treated; CR1: first haematological complete response; RFS: relapse-free survival; SoC: standard of care.

Figure 26. Schoenfeld residual plots for RFS for patients receiving blinatumomab and SoC from propensity matched analysis of BLAST and historical control study, ATT weights

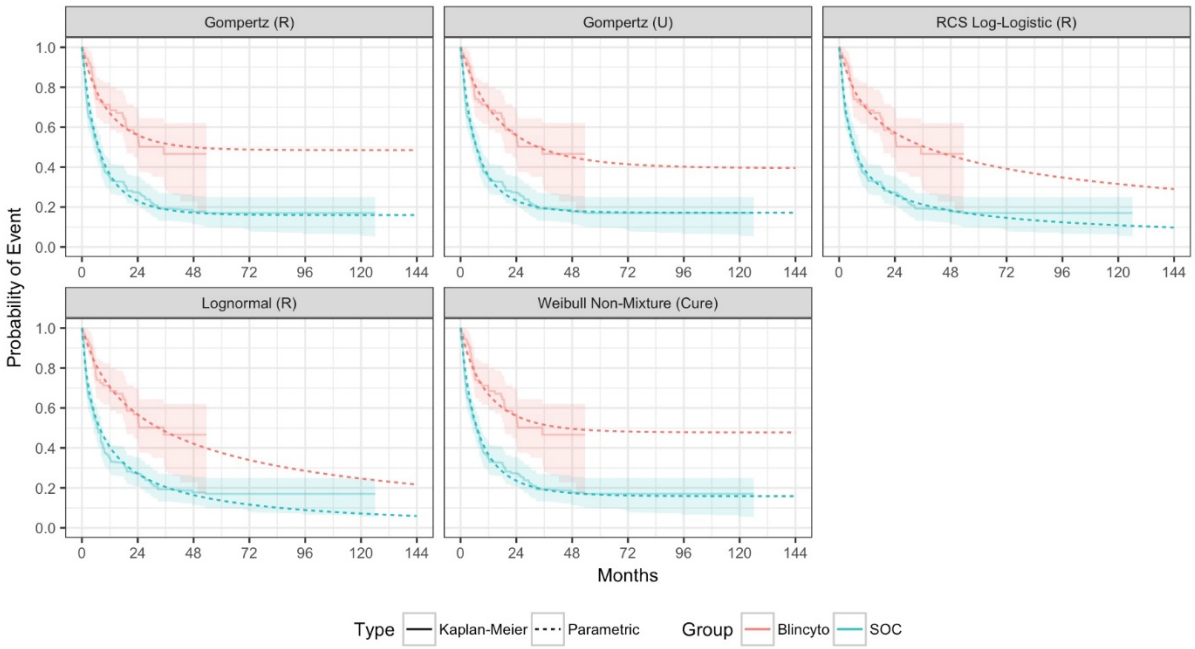


Abbreviations: ATT: average treatment effect on the treated; CR1: first haematological complete response; RFS: relapse-free survival; HR: hazard ratio.

Survival probabilities up to 12 years are shown in Figure 27. Visual fit was best for the unrestricted Gompertz distribution, which showed no systematic bias in its projections for either arm. Projections for the restricted Gompertz and Weibull non-mixture cure model both

overestimated RFS for blinatumomab at the tail. The restricted RCS log-logistic and restricted lognormal underestimated RFS for SoC beyond year four, with the latter projecting RFS at ten years at the lower bound of the 95% CI of the historical comparator RFS K-M.

Figure 27. Survival probabilities to 12 years for five best fitting parametric survival distributions fit to RFS for patients receiving blinatumomab and SoC from propensity matched analysis



Abbreviations: ATT: average treatment effect on the treated; CR1: first haematological complete response; RFS: relapse-free survival; R: restricted; U: unrestricted; SoC: standard of care.

Estimates of the cure fraction for the top models for RFS are shown in Table 36. The restricted and unrestricted Gompertz distributions, despite not being parameterised as cure models, projected plateaus in RFS for both the SoC and blinatumomab arms.

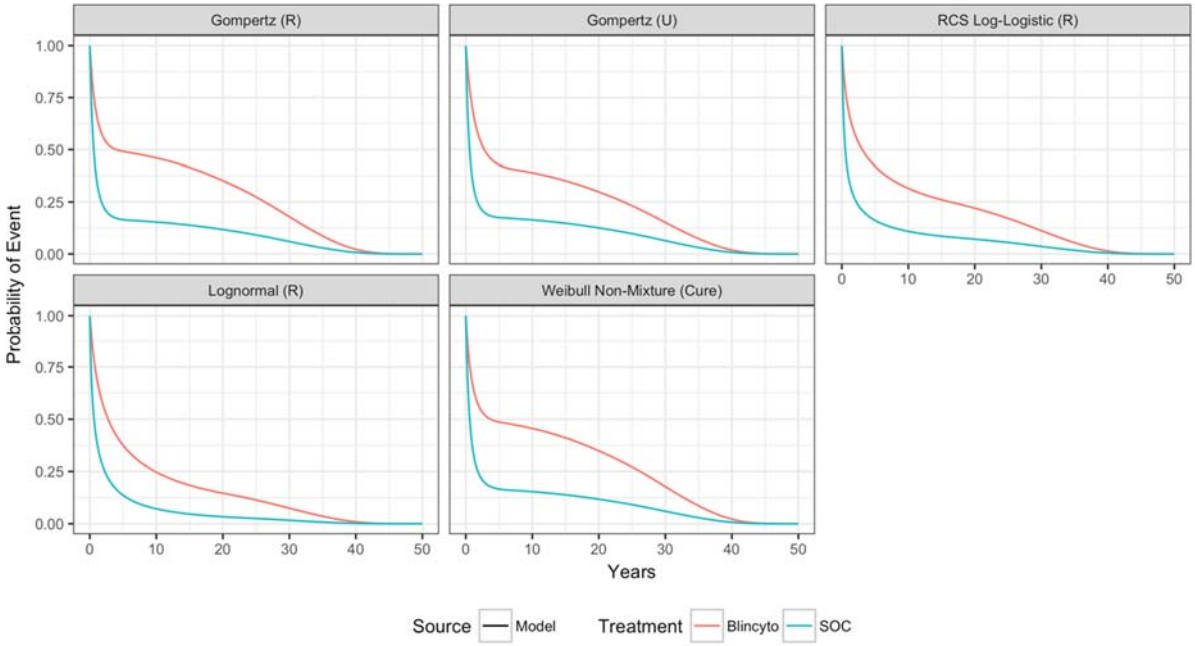
Table 36. Estimated cure fractions for five best fitting parametric survival distributions fit to RFS for patients receiving blinatumomab and SoC from propensity matched analysis of BLAST and historical control study, ATT weights

Distribution	Type	Blinatumomab	SoC
Gompertz	Restricted	48.5%	16.0%
Gompertz	Unrestricted	39.5%	17.2%
RCS Log-Logistic	Restricted	0%	0%
Lognormal	Restricted	0%	0%
Weibull Non-Mixture	Cure	47.8%	15.8%

Abbreviations: ATT: average treatment effect on the treated; CR1: first haematological complete response; RFS: relapse-free survival; SoC: standard of care; RCS: restricted cubic spline.

Model projections for each of the top five RFS distributions, incorporating age- and sex-matched adjusted general population mortality rates as a floor for the hazards, are shown in Figure 28.

Figure 28. Model projections for top five best fitting parametric survival distributions fit to RFS for patients receiving blinatumomab and SoC from propensity matched analysis of BLAST and historical control study, ATT weights



Abbreviations: ATT: average treatment effect on the treated; CR1: first haematological complete response; RFS: relapse-free survival; R: restricted; U: unrestricted; SoC: standard of care.

External validity of model projections was tested versus the shape of the projections from Berry et al. Projections based on the restricted Gompertz, unrestricted Gompertz, and Weibull non-mixture cure models were considered to have high external validity due to the similar shapes of the RFS curves as well as the similar difference in RFS at ten years. Projections based on the restricted lognormal and restricted RCS log-logistic were deemed to have lower external validity due to the overestimation of the hazard rates at 10 years, particularly for blinatumomab.

Criteria used in the selection of RFS are shown in Table 37. Only the top five best fitting RFS distributions, according to BIC, were considered. The unrestricted Gompertz distribution was chosen for use in the base-case due to its good statistical fit, visual fit, and external validity. Projections of RFS based on the unrestricted Gompertz for blinatumomab were potentially underestimated compared to Berry et al. while projections for SoC were overestimated. These projections may therefore represent a conservative projection of benefit.

Table 37. Selection criteria, ATT-weighted CR1 RFS

Distribution	Model Specification	Δ BIC	Treatment Effect	Visual Fit	External Validity	Comments
Gompertz	Restricted	--	Moderate	Moderate	Good	Counterfactual plots suggest proportional hazards may overestimate long-term benefit of blinatumomab.
Gompertz	**Unrestricted**	**3.53**	**--**	**Good**	**Good**	**Good visual fit, statistical fit, and external validity.**

RCS Log-Logistic	Restricted	3.6	Good	Moderate	Poor	Proportional odds model. Underestimates benefit of blinatumomab relative to external data.
Lognormal	Restricted	5.14	Good	Poor	Poor	Accelerated failure time model. Poor visual fit, underestimates benefit of blinatumomab relative to external data.
Weibull Non-Mixture	Cure	5.24	Moderate	Moderate	Good	Treatment effect parameterised as a cure model, but also follows proportional hazards. Counterfactual plots suggest proportional hazards may overestimate long-term benefit of blinatumomab.

Footnotes: Bolded distribution selected as base-case.
Abbreviations: ATT: average treatment effect on the treated; CR1: first haematological complete response; RFS: relapse-free survival; BIC: Bayesian information criterion; RCS: restricted cubic spline.

As noted above, calculation of the costs of allogeneic HSCT and salvage therapy requires estimates of the proportion of RFS events that are deaths. For the base-case, these estimates were based on percent of RFS events that were deaths in the propensity-matched analysis of patients in BLAST and the historical control based on the ATT weights. As shown in Table 38, based on the ATT IPTWs, the percent of RFS events that were deaths was 47.1% for blinatumomab and 8.5% for SoC. The relatively high proportion of RFS events that were deaths for blinatumomab likely reflects two factors. First, more patients underwent HSCT in BLAST, and a notable proportion of patients undergoing transplant received transplants from mismatched donors which require intensive immune suppressive medication to prevent host rejection by the graft which leads to increased risk of severe, often deadly, infections. Second, capture of relapses after transplant may have been incomplete in BLAST. It was not feasible to address the extent of the underreporting of relapse in BLAST. Accordingly, for the base-case, the estimate from BLAST was used. A scenario analysis was conducted assuming the proportion of RFS that were deaths was only 20%.

Table 38. Distribution of RFS events for patients in the propensity-matched analysis of BLAST and the historical control study

RFS Events	BLAST (Blinatumomab)		Historical Control (SoC)	
	N	%	N	%
Unweighted				
Death	16	47.1%	14	10.7%
Relapse	18	52.9%	117	89.3%

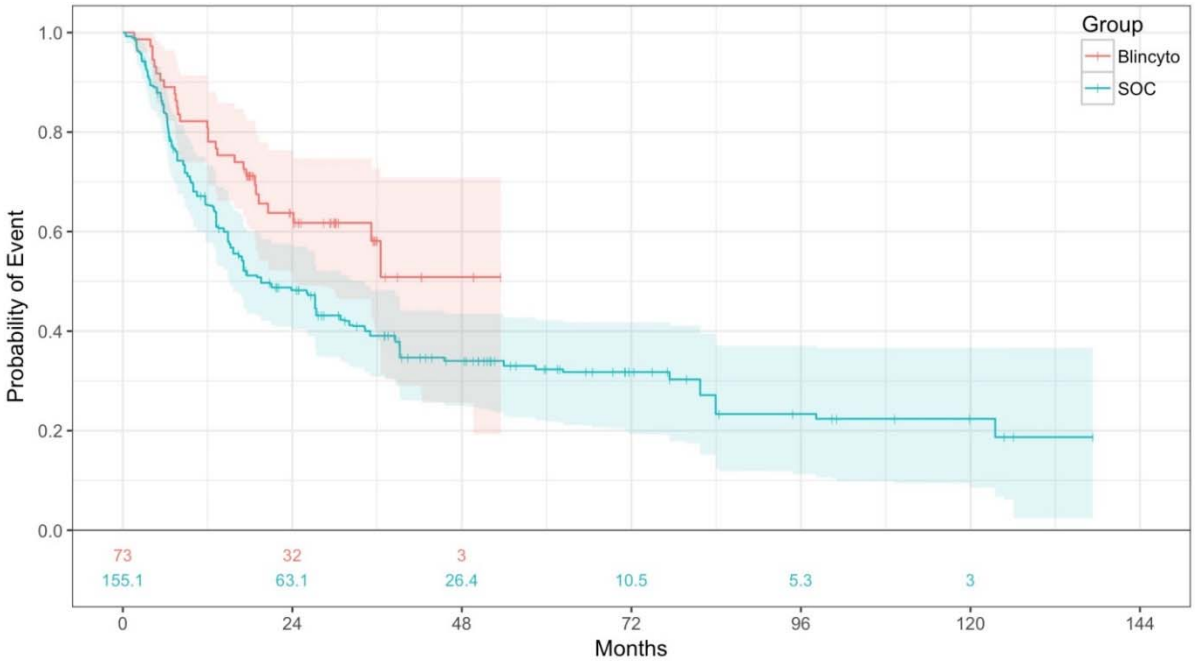
Total	34	100.0%	131	100.0%
ATT-IPTW				
Death	16	47.1%	10.4	8.5%
Relapse	18	52.9%	112	91.6%
Total	34	100.0%	122.3	100.0%
ATE-IPTW				
Death	13.8	40.2%	13	10.1%
Relapse	20.5	59.8%	115.6	90.0%
Total	34.3	100.0%	128.5	100.0%

Abbreviations: RFS: relapse-free survival; SoC: standard of care; ATT: average treated effect on the treated; IPTW: inverse probability of treatment weighting; ATE: average treated effect.

Overall Survival

K-M estimates of ATT-weighted OS in the CR1 subgroup are shown in Figure 22. Clear separation was observed between blinatumomab and SoC throughout the follow-up of BLAST. As with RFS, a strong pattern of decreasing hazards was observed in both arms.

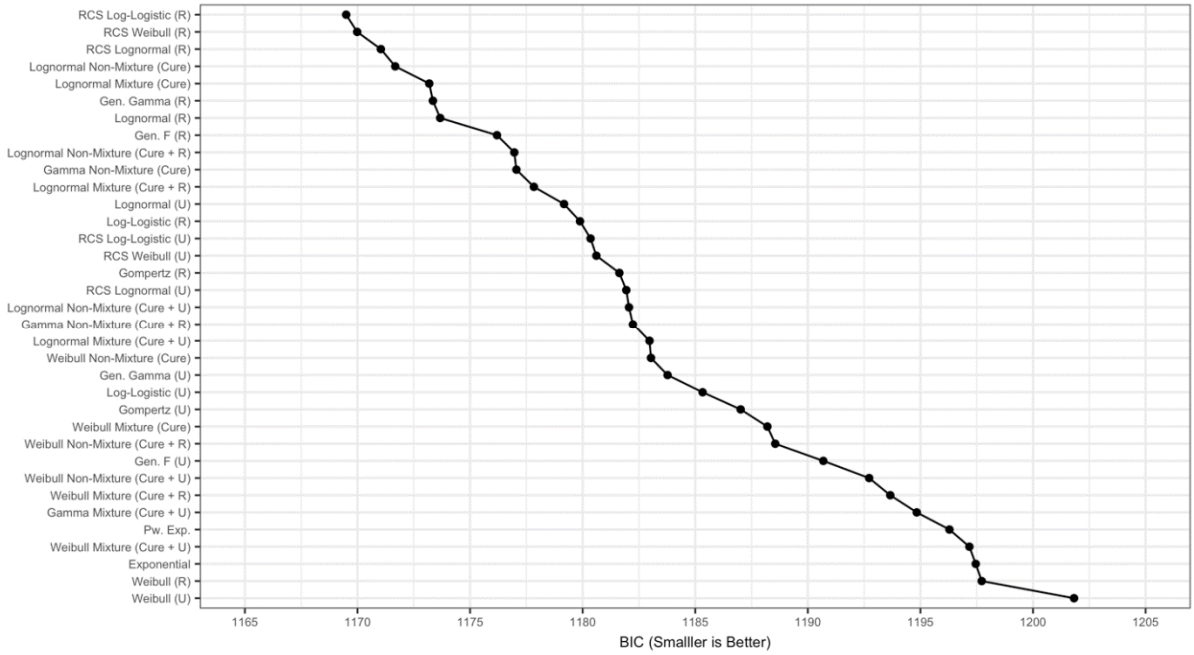
Figure 29. K-M estimates of OS for patients receiving blinatumomab and SoC from propensity matched analysis of BLAST and historical control study, ATT weights



Abbreviations: ATT: average treated effect on the treated; CR1: first haematological complete response; OS: overall survival; SoC: standard of care.

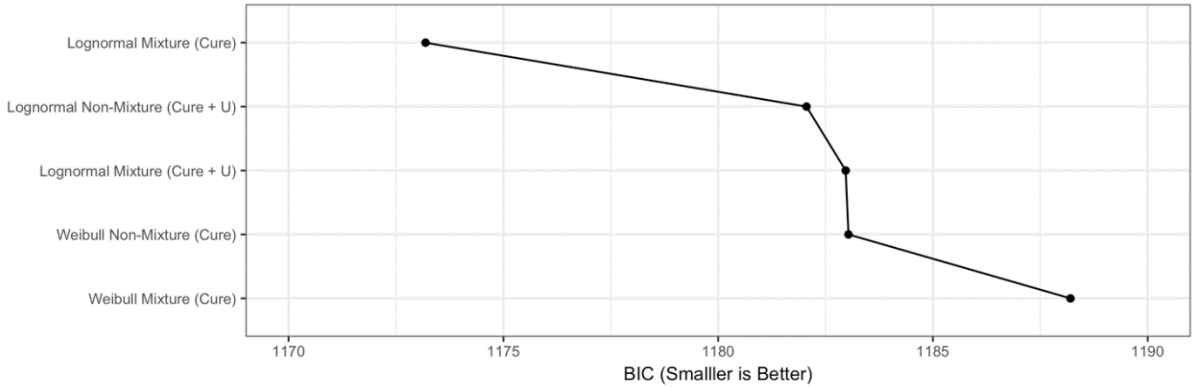
Statistical fit for all fitted distribution, according to BIC, is shown in Figure 30. In order to focus on the best fitting distributions and maintain internal consistency with the selected base-case RFS distributions, only the top five best-fitting distributions which did not cross (i.e. OS ≥ RFS throughout the model projection) the RFS unrestricted Gompertz distribution were considered. Additional information on all fitted OS distributions can be found in Appendix P. Statistical fit for the top five qualifying distributions is shown in Figure 31.

Figure 30. Fit statistics all parametric distributions fit to OS for patients receiving blinatumomab and SoC from propensity matched analysis of BLAST and historical control study, ATT weights



Abbreviations: ATT: average treatment effect on the treated; CR1: first haematological complete response; OS: overall survival; R: restricted; U: unrestricted; RCS: restricted cubic spline; BIC: Bayesian information criterion.

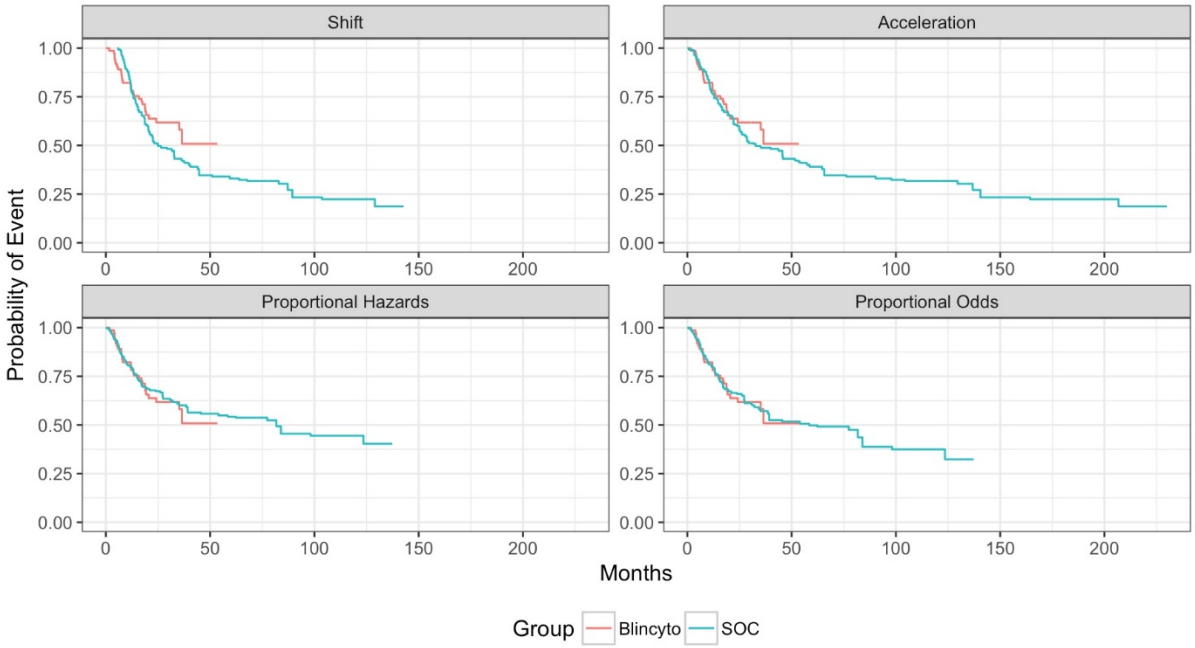
Figure 31. Fit statistics for top five qualifying fitted distributions, ATT-weighted CR1 OS



Abbreviations: ATT: average treatment effect on the treated; CR1: first haematological complete response; OS: overall survival; R: restricted; U: unrestricted; BIC: Bayesian information criterion.

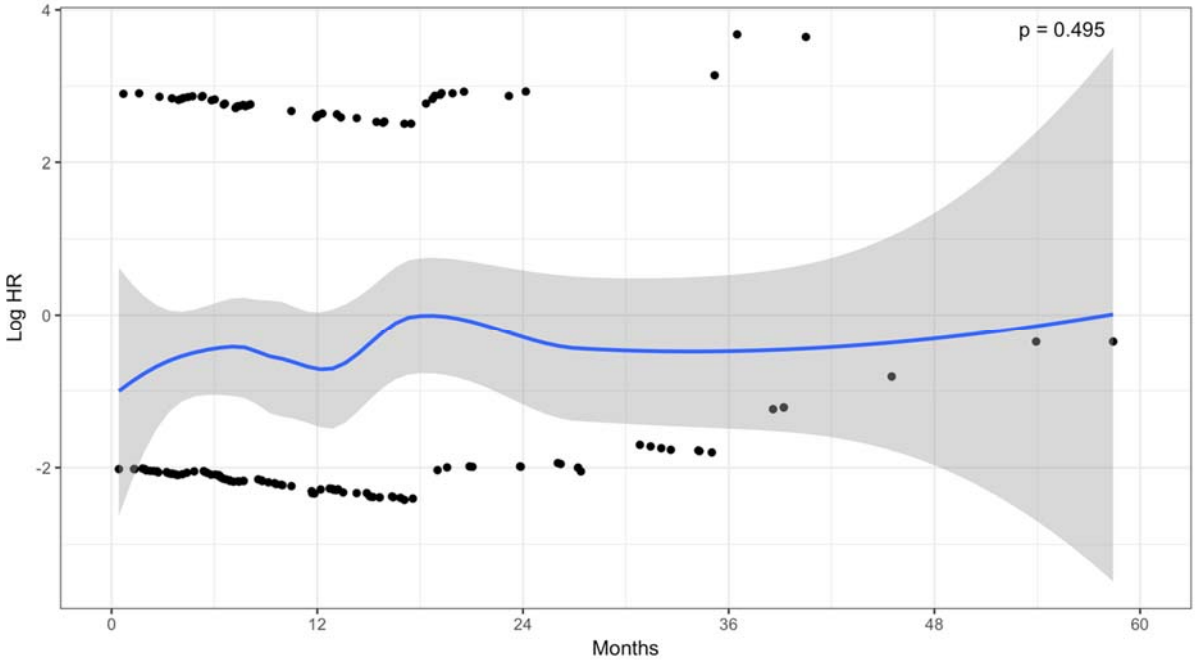
Treatment effect counterfactual plots are shown in Figure 32. OS from BLAST and the historical comparator study were well represented by proportional odds, accelerated failure time, and proportional hazards, with no systematic bias observed. A fixed shift in RFS was ruled out due systematic bias throughout the trial follow-up. No significant deviation from proportional hazards was identified in the Schoenfeld residuals (Figure 33).

Figure 32. Treatment effect counterfactual plots for OS for patients receiving blinatumomab and SoC from propensity matched analysis of BLAST and historical control study, ATT weights



Abbreviations: ATT: average treatment effect on the treated; CR1: first haematological complete response; OS: overall survival; SoC: standard of care.

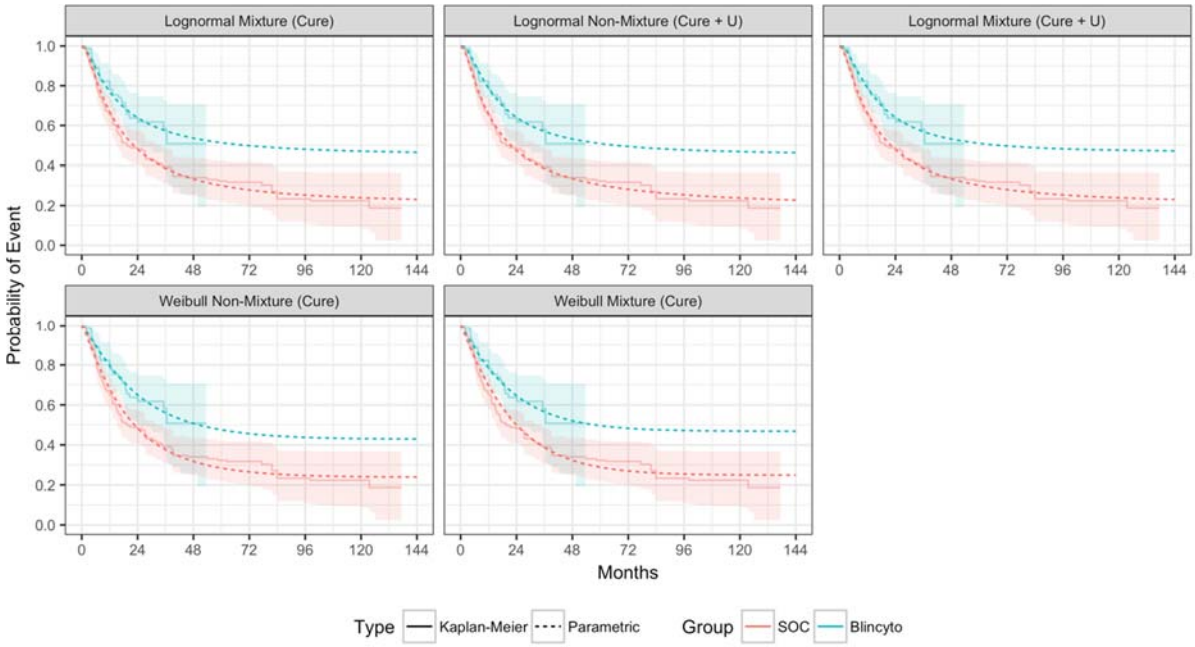
Figure 33. Schoenfeld residuals for OS for patients receiving blinatumomab and SoC from propensity matched analysis of BLAST and historical control study, ATT weights



Abbreviations: ATT: average treatment effect on the treated; CR1: first haematological complete response; OS: overall survival; HR: hazard ratio.

Survival probabilities for each of the top five fitted distributions, compared with the corresponding K-M curves, are shown to 12 years in Figure 34. Visual fit was acceptable for all of the top five distributions.

Figure 34. Survival probabilities to 12 years for five best fitting parametric distributions fit to OS for patients receiving blinatumomab and SoC from propensity matched analysis



Abbreviations: ATT: average treatment effect on the treated; CR1: first haematological complete response; OS: overall survival; R: restricted; U: unrestricted; SoC: standard of care.

Estimates of the cure fraction for the top models for OS are shown in Table 39. Since only models consistent with the RFS restricted Gompertz – which projected a plateau in RFS for both arms – were considered, all remaining models projected a plateau in OS for both SoC and blinatumomab.

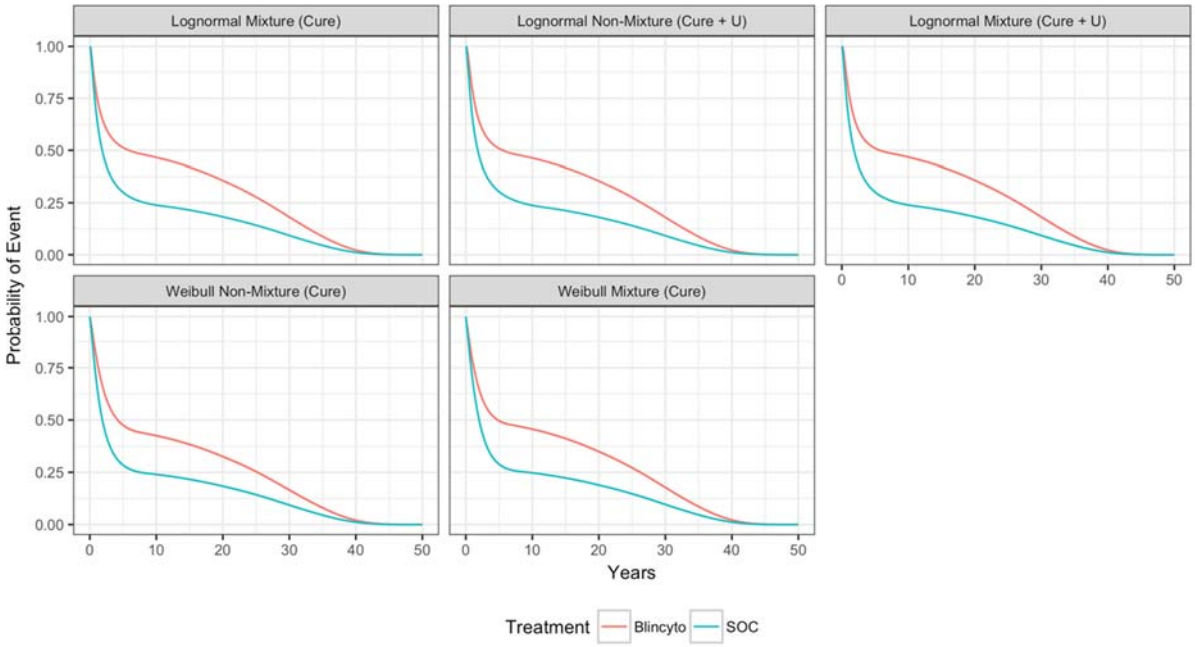
Table 39. Estimated cure fractions for five best fitting parametric distributions fit to OS for patients receiving blinatumomab and SoC from propensity matched analysis of BLAST and historical control study, ATT weights

Distribution	Type	Blinatumomab	SoC
Lognormal Mixture	Cure	45.3%	21.3%
Lognormal Non-Mixture	Cure + Unrestricted	45.3%	19.3%
Lognormal Mixture	Cure + Unrestricted	46.6%	21.0%
Weibull Non-Mixture	Cure	42.8%	23.8%
Weibull Mixture	Cure	46.8%	24.9%

Abbreviations: ATT: average treatment effect on the treated; CR1: first haematological complete response; OS: overall survival; SoC: standard of care.

Model projections for each of the top five OS distributions, with background mortality incorporated, are shown in Figure 35.

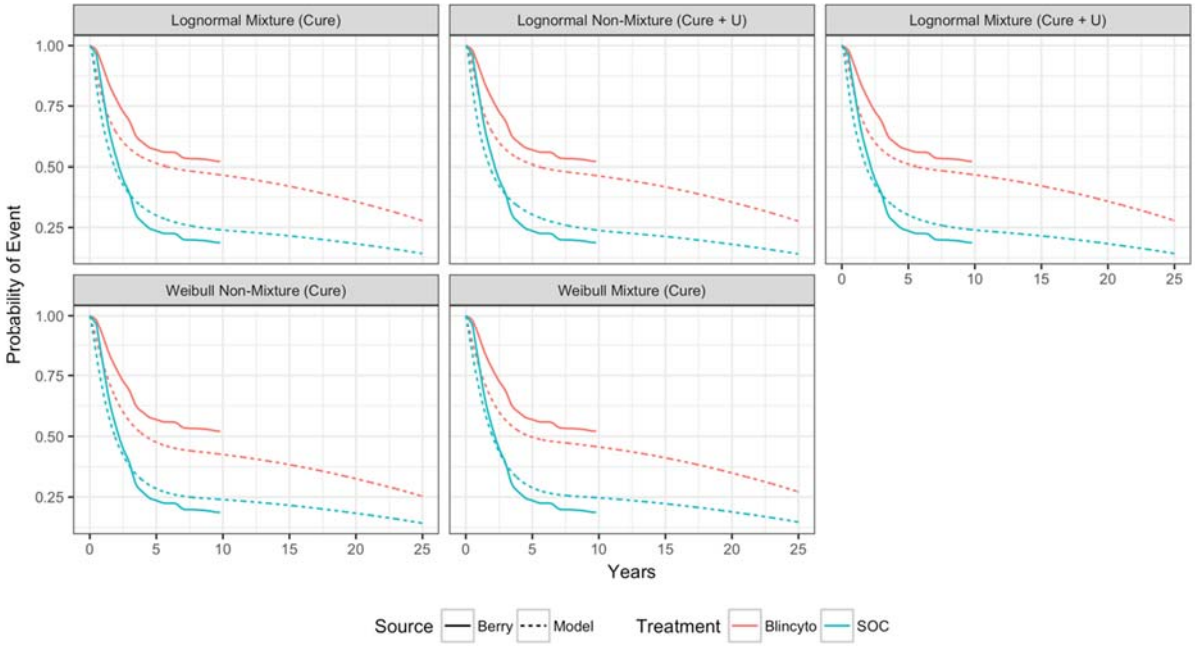
Figure 35. Model projections for five best fitting parametric distributions fit to OS for patients receiving blinatumomab and SoC from propensity matched analysis



Abbreviations: ATT: average treatment effect on the treated; CR1: first haematological complete response; OS: overall survival; R: restricted; U: unrestricted; SoC: standard of care.

External validity of model projections for the top five OS models were tested using projections based on Berry et al., as shown in Figure 36. For blinatumomab, all five models overestimated hazards relative to the external data but yielded a similar shape with a plateau slightly below that of the external data. For SoC, hazard rates were overestimated initially but underestimated afterwards resulting in the curves crossing and a slightly higher plateau than that of the external data. All model projections were therefore considered to have moderate external validity.

Figure 36. MRD response adjusted OS from Berry et al. compared with model projections for five best fitting parametric distributions fit to OS for patients receiving blinatumomab and SoC from propensity matched analysis



Abbreviations: MRD: minimal residual disease; EFS: event-free survival; ATT: average treatment effect on the treated; CR1: first haematological complete response; OS: overall survival; R: restricted; U: unrestricted; SoC: standard of care.

Criteria used in the selection of the base-case OS distributions are shown in Table 40. The lognormal mixture cure model was selected for use in the base-case due to its much better statistical fit than the other distributions considered. The difference in BIC between the lognormal mixture cure model and the next best fitting model was 8.87, within the range of 6-10 generally considered strong evidence.[92] Projections using this model may be conservative as they project a smaller gain in survival than that implied by the external data, and a decrement after approximately seven months in PRS.

Table 40. Selection criteria, ATT-weighted CR1 OS

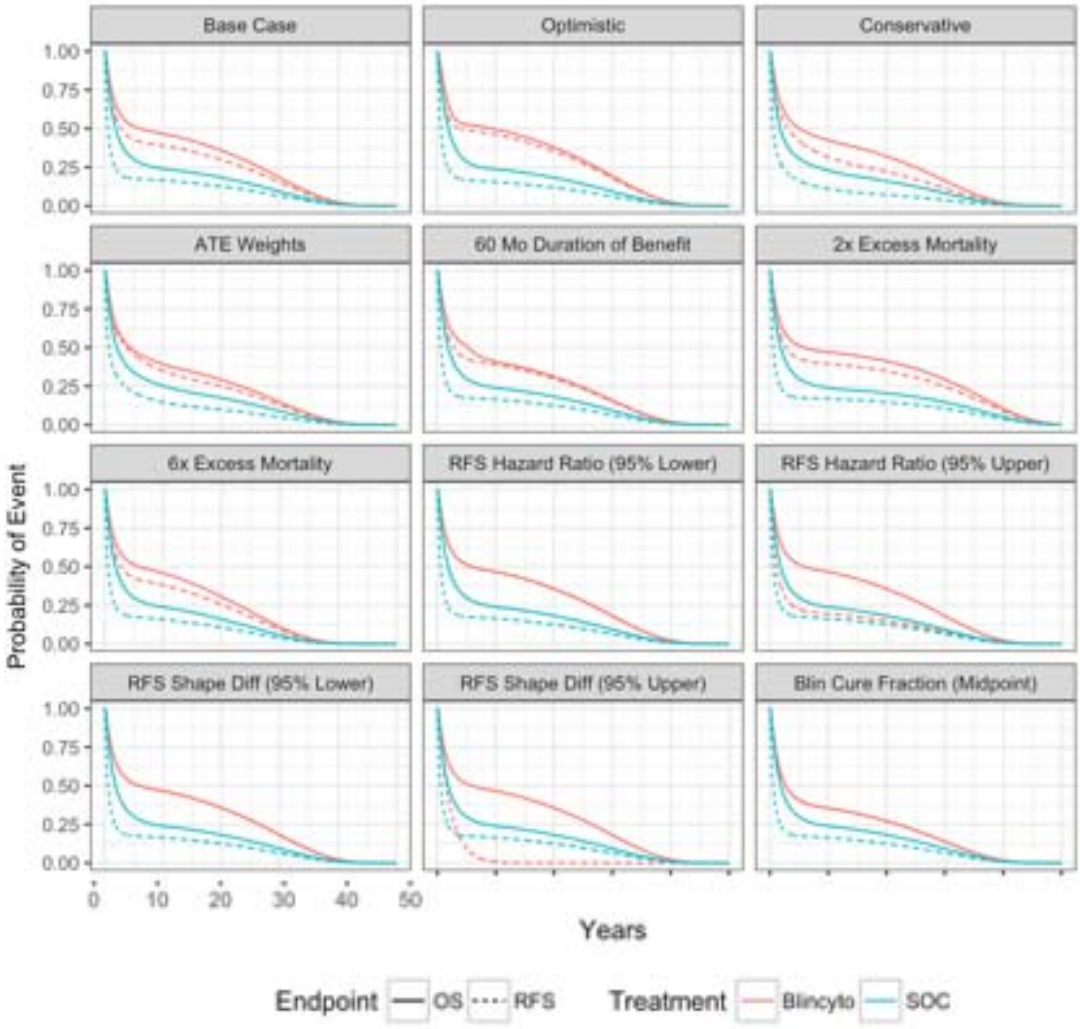
Distribution	Model Specification	Δ BIC	Treatment Effect	Visual Fit	External Validity	Δ PRS (Years)	Comments
Lognormal Mixture	Cure	--	--	Good	Moderate	-0.70	Best-fitting distribution among those consistent with base-case RFS. Large difference in BIC versus next best-fitting distribution.
Lognormal Non-Mixture	Cure + Unrestricted	8.87	--	Good	Moderate	-0.69	Poor statistical fit.
Lognormal Mixture	Cure + Unrestricted	9.78	--	Good	Moderate	-0.65	Poor statistical fit.
Weibull Non-Mixture	Cure	9.85	Good	Good	Moderate	-1.58	Poor statistical fit. Treatment effect counterfactual plots are supportive of proportional hazards. Large difference in PRS.
Weibull Mixture	Cure	15.02	--	Good	Moderate	-1.11	Poor statistical fit. Large difference in PRS.

Footnotes: Bolded distribution selected as base-case.
Abbreviations: ATT: average treatment effect on the treated; CR1: first haematological complete response; OS: overall survival; BIC: Bayesian information criterion; PRS: post-relapse survival.

Sensitivity Analyses

In addition to the selected base-case distributions, alternate RFS and OS distributions were used in several sensitivity analyses to test the sensitivity of model results to both structural assumptions and parameter uncertainty. Alternate distributions of RFS and OS used in each sensitivity analysis are shown in Figure 37.

Figure 37. Model projections of RFS and OS used in sensitivity analyses



Footnotes: Panels for OS blinatumomab cure fraction 95% lower and upper bounds are generated by varying the cure fraction only and do not account for correlation of parameters for each distribution.
Abbreviations: RFS: relapse-free survival; OS: overall survival; ATE: average treatment effect; Mo: month; Blin: blinatumomab; SoC: standard of care.

Alternate parametric forms for RFS and OS were considered in the “more favourable” scenario and “less favourable” scenarios. For each scenario, an alternate RFS distribution was selected to anchor model projections followed by reapplying selection criteria to obtain an internally consistent projection of OS.

In the first alternative curve, the more favourable restricted Gompertz model was used to model RFS. While this model assumes proportional hazards, which appeared to overestimate benefit towards the end of follow-up of BLAST, it had the best statistical fit of any distribution considered and appeared to somewhat more accurately reflect the magnitude of benefit as predicted by the

external data on RFS by MRD response. Selection criteria used to identify the OS distribution used in the more favourable scenario are shown in Table 41. Only two fitted OS distributions produced projections which did not cross the RFS restricted Gompertz and could therefore be considered for use in this scenario. Of these two distributions, the Weibull non-mixture cure plus unrestricted model was selected based on better statistical fit, as the two distributions were graded near identically in all other criteria.

Table 41. Selection criteria for OS distribution used in more favourable scenario

Distribution	Model Specification	Δ BIC	Treatment Effect	Visual Fit	External Validity	Δ PRS (Years)	Comments
Weibull Non-Mixture	Cure + Unrestricted	--	--	Good	Moderate	-1.83	Best statistical fit of qualifying distributions.
Weibull Mixture	Cure + Unrestricted	4.45	--	Good	Moderate	-1.95	Worse statistical fit, no other distinguishing features.

Footnotes: Bolded distribution selected as base-case.
Abbreviations: OS: overall survival; BIC: Bayesian information criterion; PRS: post-relapse survival.

The RCS log-logistic distribution was selected to represent RFS in the less favourable scenario, as it had the third best statistical fit; some supportive evidence in favour of proportional odds based on the treatment effect counterfactual plots, and appeared to represent a plausible lower bound on the benefit of treatment with blinatumomab. Selection criteria used to identify the OS distribution used in the less favourable scenario are shown in Table 42 below. Of the top five best fitting OS distributions among those internally consistent with the RFS restricted RCS log-logistic distribution, the restricted RCS Weibull distribution was selected to model OS due to its combination of statistical fit, supportive evidence in favour of proportional hazards, external validity, and small difference in PRS.

Table 42. Selection criteria for OS distribution used in less favourable scenario

Distribution	Model Specification	Δ BIC	Treatment Effect	Visual Fit	External Validity	Δ PRS (Years)	Comments
RCS Log-Logistic	Restricted	--	Good	Good	Poor	-1.76	Best-fitting OS distribution and supportive evidence for proportional odds, but poor external validity and projects large difference in PRS.
RCS Weibull	**Restricted**	**0.49**	**Good**	**Good**	**Moderate**	**-0.61**	**Supportive evidence in favour of proportional hazards, moderate external validity, no meaningful difference in BIC versus best fitting distribution.**
RCS Lognormal	Restricted	1.54	--	Good	Poor	-1.99	Poor external validity and projects large difference in PRS.
Lognormal Non-Mixture	Cure	2.18	Good	Good	Moderate	-1.55	Supportive evidence in favour of proportional hazards, moderate external validity, and large difference in PRS.
Lognormal Mixture	Cure	3.69	--	Good	Moderate	-0.14	Moderate external validity. Small difference in PRS.

Footnotes: Bolded distribution was selected for use in sensitivity analysis.
Abbreviations: OS: overall survival; BIC: Bayesian information criterion; PRS: post-relapse survival; RCS: restricted cubic spline.

To evaluate the sensitivity of model results to the weighting scheme used, a scenario in which the ATE weights were used was performed. RFS and OS distributions fitted using the ATE weights were selected based on the same criteria used in the base-case. Selection criteria used to identify the RFS distribution for the ATE-weighted scenario are shown in Table 43. Of the top five distributions, only the restricted Gompertz had good external validity, but was not selected due to the lack of supporting evidence to justify assuming proportional hazards. The restricted RCS log-logistic distribution was selected to model RFS as it had the greatest external validity among the remaining distributions. For OS, the restricted RCS Weibull based on its combination of statistical fit, external validity, internal consistency with the selected RFS distribution. The selection criteria used for the selection of the OS distribution are outlined in Table 44. Additional information on all fitted RFS and OS distributions based on the ATE-weighted sample can be found in Appendix P.

Table 43. Selection Criteria for RFS Distribution Used in ATE-Weighted Scenario

Distribution	Model Specification	Δ BIC	Treatment Effect	Visual Fit	External Validity	Comments
Lognormal	Restricted	--	Good	Moderate	Poor	Proportional hazards model. Counterfactual plots suggest proportional hazards may overestimate long-term benefit of blinatumomab.
Generalized Gamma	Restricted	0.45	Good	Good	Poor	Good visual fit, statistical fit, and external validity.
RCS Log-Logistic	**Restricted**	**1.23**	**Good**	**Moderate**	**Moderate**	**Proportional odds model. Underestimates benefit of blinatumomab relative to external data.**
RCS Lognormal	Restricted	1.39	--	Poor	Poor	Accelerated failure time model. Poor visual fit, underestimates benefit of blinatumomab relative to external data.
Gompertz	Restricted	2.26	Poor	Moderate	Good	Treatment effect parameterised as a cure model, but follows proportional hazards. Counterfactual plots suggest proportional hazards may overestimate long-term benefit of blinatumomab.

Footnotes: Bolded distribution was selected for use in sensitivity analysis.
Abbreviations: RFS: relapse-free survival; ATE: average treatment effect; BIC: Bayesian information criterion; RCS: restricted cubic spline.

Table 44. Selection Criteria for OS Distribution Used in ATE-Weighted Scenario

Distribution	Model Specification	Δ BIC	Treatment Effect	Visual Fit	External Validity	Δ PRS (Years)	Comments
RCS Log-Logistic	Restricted	--	Good	Good	Poor	-2.38	Large decrease in post-relapse survival, poor external validity.
RCS Weibull	**Restricted**	**1.56**	**Moderate**	**Good**	**Moderate**	**-1.90**	**Smallest difference in PRS.**
Lognormal Non-Mixture	Cure	3.86	Moderate	Good	Moderate	-2.11	Proportional odds model. Underestimates benefit of

							blinatumomab relative to external data.
Lognormal Mixture	Cure	6.11	--	Good	Poor	-2.16	Poor statistical fit, large decrease in PRS, poor external validity.
Gamma Non-Mixture	Cure	8.61	Moderate	Good	Good	-2.13	Very poor statistical fit, large decrease in PRS.

Footnotes: Bolded distribution was selected for use in sensitivity analysis.
Abbreviations: OS: overall survival; ATE: average treatment effect; BIC: Bayesian information criterion; PRS: post-relapse survival; RCS: restricted cubic spline.

Scenarios were also explored to test the sensitivity of model results to limiting the duration of benefit for RFS and OS to five years, as well as to varying background mortality between two and six times the age- and gender-matched rate. Finally, to capture the effect of parameter uncertainty on RFS, sensitivity analyses were performed to vary the difference in shape and scale between blinatumomab and SoC over their respective 95% confidence intervals.

B.3.4 *Measurement and valuation of health effects*

B.3.4.1 Health-related quality-of-life data from clinical trials

Utility values for pre-relapse states were based on EQ-5D utility values for the patients in BLAST who were included in the PAS of the propensity-matched comparison of patients in BLAST and the historical control study (N=73). EQ-5D utility values were calculated using UK tariffs. Utility values used in the model were estimated using a generalised linear model/generalised estimating equations (GLM/GEE) regression model with EQ-5D utility values as the dependent variable and covariates for baseline utility value, a patient-level indicator variable of MRD response during cycles 1 or 2, a time-dependent indicator variable for on versus off treatment, and a time-dependent indicator variable for death within 6 months. Patients without any follow-up utility assessments were excluded, as were those without baseline utility values. There were eight assessments conducted on or after relapse, most of which were evaluated on the day of relapse. As these observations are not likely representative of quality of life during the entire post-relapse period, utility assessments on or after relapse also were excluded from the analysis. A total of 63 patients were included in the final regression model used for the base-case estimates.[*]

The number of utility assessments per strata defined on the covariate values is reported in Table 45.

Table 45. Number of utility assessments in GLM/GEE models by covariate strata

Strata	N of utility assessments
Death within ≤ 6 month	183
Death within > 6 month	13
On-treatment	123
Post-treatment	73
MRD responder: No	13
MRD responder: Yes	183

Abbreviations: GLM: generalised linear model; GEE: generalised estimating equations; MRD: minimal residual disease.

Parameter estimates from the GLM/GEE model are reported in Table 46. While the coefficients for off versus on treatment and MRD response were not statistically significant, they were

[*]Of the 73 patients in BLAST in the PAS of the propensity score analysis, one patient lacked a propensity score and was excluded from analyses using ATE weights.

directionally consistent with expectations (i.e., higher utility values). These covariates were therefore retained when calculating utility values used in the model.

Table 46. Parameter estimates from GLM/GEE regression on EQ-5D utility values among CR1 patients in BLAST

Parameter	Value	SE	95% CI-L	95% CI-U	P-value
Intercept	0.3531	0.0918	0.1732	0.5329	<.001
Baseline utility	0.5427	0.0832	0.3797	0.7058	<.001
Off- versus on-treatment relapse free	0.0105	0.0168	-0.0225	0.0434	0.5347
MRD response versus no MRD response	0.0474	0.0469	-0.0446	0.1394	0.3125
Death within ≤6 month versus death within >6 month	-0.1291	0.0346	-0.1970	-0.0613	<.001

Abbreviations: GLM: generalised linear model; GEE: generalised estimating equations; EQ-5D: EuroQol Five Dimensions Questionnaire; CR1; first haematological complete remission; SE: standard error; CI-L: lower bound of the confidence interval; CI-U: upper bound of the confidence interval; MRD: minimal residual disease.

Utility values during RFS were calculated by treatment group and cycle using this regression equation, the mean baseline utility for CR1 patients in BLAST (0.809), the estimated proportion of patients receiving blinatumomab treatment, and the estimated proportion of patients with MRD response. For patients receiving blinatumomab, the proportion with MRD response was based on the percent of patients with MRD response at the end of cycle 2 in BLAST (83.6%). As noted above, for patients receiving SoC maintenance therapy, the proportion of patients who might achieve a delayed MRD response is unknown, as this information was not examined in the historical control study, and has not been reported in the literature. Discussions with clinical experts indicate that this proportion is no greater than 10%. It was therefore assumed that 8% of patients receiving SoC maintenance therapy would achieve a delayed MRD response. Patients achieving MRD response were assumed to achieve response at the end of the first cycle and remain in response until death or relapse. In BLAST, the median duration of response among patients achieving response in the key secondary efficacy endpoint FAS was 17.3 months, which compares with the median RFS of 18.9 moths. The assumption that patients remain in MRD response until relapse is therefore consistent with these data. For patients receiving blinatumomab, the proportion of patients on treatment was estimated using exposure data from BLAST as described in Section B.3.5 below.

Because post-relapse utility assessments in BLAST were limited and not likely representative of utility during the entire post-relapse period, post-relapse utility estimates were not obtained from BLAST. Rather, they were based on estimated utility values for patients receiving SoC salvage chemotherapy in the TOWER trial of blinatumomab in Ph- R/R B-precursor cell ALL who were matched to patients who relapsed in BLAST. Utility values in the TOWER study were based on a mapping from the EORTC QLQ-C30 to EQ-5D utility values (UK tariffs) using a published algorithm by Longworth.[93]

Relapsed patients in the CR1 population of BLAST can be considered similar to patients in TOWER without prior salvage (S0) who were not refractory at baseline. Of the 34 relapsed patients in the CR1 population of BLAST, 13 patients relapsed more than 12 months after therapy initiation. Since the TOWER inclusion criteria specify that patients with no prior salvage therapy must have relapsed within 12 months of remission, these 13 BLAST patients are not represented in the TOWER study and were excluded from the analysis. The remaining patients

in TOWER were matched to the relapsed patients in BLAST based on their health state: i.e., CR1/CR2 (BLAST) or S0/S1 (S1 refers to the patients with one or more prior lines of salvage therapy) (TOWER), age, and their receipt of allogeneic HSCT (at baseline among TOWER patients and prior to relapse among BLAST patients). A logistic regression model was estimated predicting the probability of being in BLAST versus TOWER. Using the estimated predicted probability of being in BLAST (versus TOWER), ATT weights were calculated for TOWER patients. Using these weights to match the patients in TOWER to those in BLAST, the mean EQ-5D utility value for patients in the S0 subgroup of the SoC chemotherapy arm of TOWER was estimated to be 0.692. This value was used to represent the mean utility value for patients who relapsed in the model.

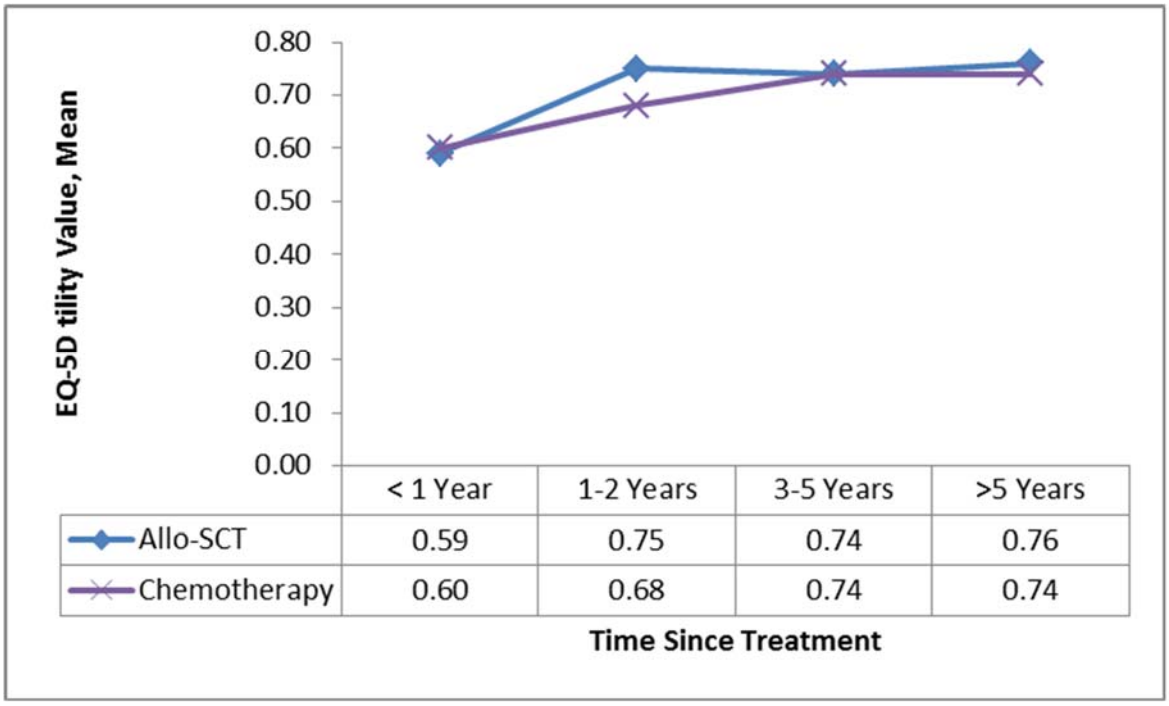
Utility values used in the model are summarised in Table 47. As noted above, there were 13 patients who relapsed after 12 months of remission in BLAST, who could not be matched to patients in TOWER. If post-relapse utility values for patients with late relapse are higher than those for patients with early relapse, then the estimates of post-relapse utility derived from TOWER might be downwardly biased.

As noted in Section B.3.2, it was assumed in the base-case that utility values for patients remaining alive after five years return to age- and sex-matched general population norms adjusted for long-term effects on HRQoL due to exposure to radiotherapy, chemotherapy, and HSCT. This decrement was assumed to be 0.02, based on one-half the difference between the average utility value for blinatumomab patients in the RFS state, off therapy, and with MRD response (0.842) versus the age and sex- weighted mean population norm utility value for patients between the ages of 35 and 55 (0.877).

Studies of HRQoL after HSCT suggest that there is substantial short-term decline in HRQoL for the first year after HSCT, with small to moderate decrements in HRQoL relative to general population norms in the long-term.[94, 95] Data on EQ-5D utility values post-HSCT in patients with acute leukaemia are limited, however. A review of HRQoL in acute myeloid leukaemia (AML) patients by Korol et al. (2017) cites 3 papers reporting EQ-5D utilities in patients with AML,[96] including a survey of 524 patients with acute leukaemia in Japan by Kurosawa et al. (2016),[97] a survey of 92 AML patients in The Netherlands by Leunis et al. (2014),[98] and a study by Slovacek and colleagues of 12 adult patients with AML undergoing autologous progenitor SCT in the Czech republic (article in Czech, not discussed further).[99] The utility values reported in the survey by Kurosawa were subsequently used in a decision analysis by the same author of post-remission therapy in cytogenetically intermediate-risk AML.[100] A targeted search of the literature and of utility values in the New England Medical Center CEA registry did not identify any additional studies reporting EQ-5D utility values in patients with acute leukaemia after HSCT. Only the decision analysis by Kurosowa et al. reports utility values by time since last treatment (HSCT or chemotherapy). Utility values from the study by Kurosawa were used in the manufacturer's submission to NICE for the STA of inotuzumab in R/R B-precursor ALL.[85]

The study by Kurosawa et al. was a cross-sectional survey of 524 patients with acute leukaemia (75% AML, 25% ALL) in Japan in 2011 and 2012. Utility values were based on the EQ-5D index using the value set for Japan. In the decision analysis that used the utility values from this study, utility values were reported by last treatment received (HSCT or chemotherapy). As shown in Figure 38 for both patients who received HSCT and chemotherapy, mean utility values increase by time since treatment and plateaued within 5 years of treatment. For HSCT, the plateau was reached at 1-2 years. For chemotherapy, the plateau was reached at 3-5 years.

Figure 38. Mean EQ-5D utility values among survivors of acute leukaemia in Japan, by treatment and time since treatment reported by Kurosawa et al. (2016)



Abbreviations: EQ-5D: EuroQol 5 dimensions questionnaire; allo-SCT; allogeneic haematopoietic stem cell transplantation.

Use of the utility estimates from Kurosawa et al. directly in the model would be consistent with the approach used in the inotuzumab NICE manufacturer's submission. However, this approach may be biased due to differences in the patients included in the study by Kurosawa et al. compared with patients in BLAST, as comparisons of the utility values in the study by Kurosawa et al. with those from BLAST suggest that the utility values may not be comparable. For example, the mean utility value for patients <1 year since chemotherapy in Kurosawa et al. (0.60) is well below the mean utility values for patients in RFS and >6 months from death in BLAST (0.66 to 0.71). The mean utility value for patients <1 year since CT in Kurosawa et al. (0.60) is only slightly greater than the estimated mean utility values for post-relapse and <6 months from death used in the model based on data from TOWER (0.562). The utility values from Kurosawa et al. also are lower than the mean utility values for AML patients with prior HSCT reported by Leunis et al. (2014), which reported a mean utility value for AML patients with prior HSCT and a mean of 5.3 years since last treatment of 0.82.[98]

To avoid potential biases associated with using the utility values from Kurosawa directly, but to capture the short-term impact of HSCT on HRQoL, all patients undergoing HSCT were assumed to experience decrements in utility of 0.17, 0.01, and 0.02 in years 1, 2, and 3-5 after HSCT, respectively, based on the differences in the mean utility value at these time points versus at >5 years post HCT (0.76) reported by Kurosawa et al. The application of these decrements in utility, along with the assumption that patients in the model who survive >60 months will have utility values equal to general population norms adjusted for the long-term decrement in utility associated with chemotherapy, radiotherapy, and HSCT (assumed to be 0.02 in the base-case), will yield estimates of utility that are consistent with the general pattern of HRQoL reported in numerous studies which suggests (1) a short-term decline for the first year after HSCT and (2) small to moderate long-term decrements relative to general population norms.

Utility values used in the cost-effectiveness analysis are summarised in Table 47 and in Figure 39.

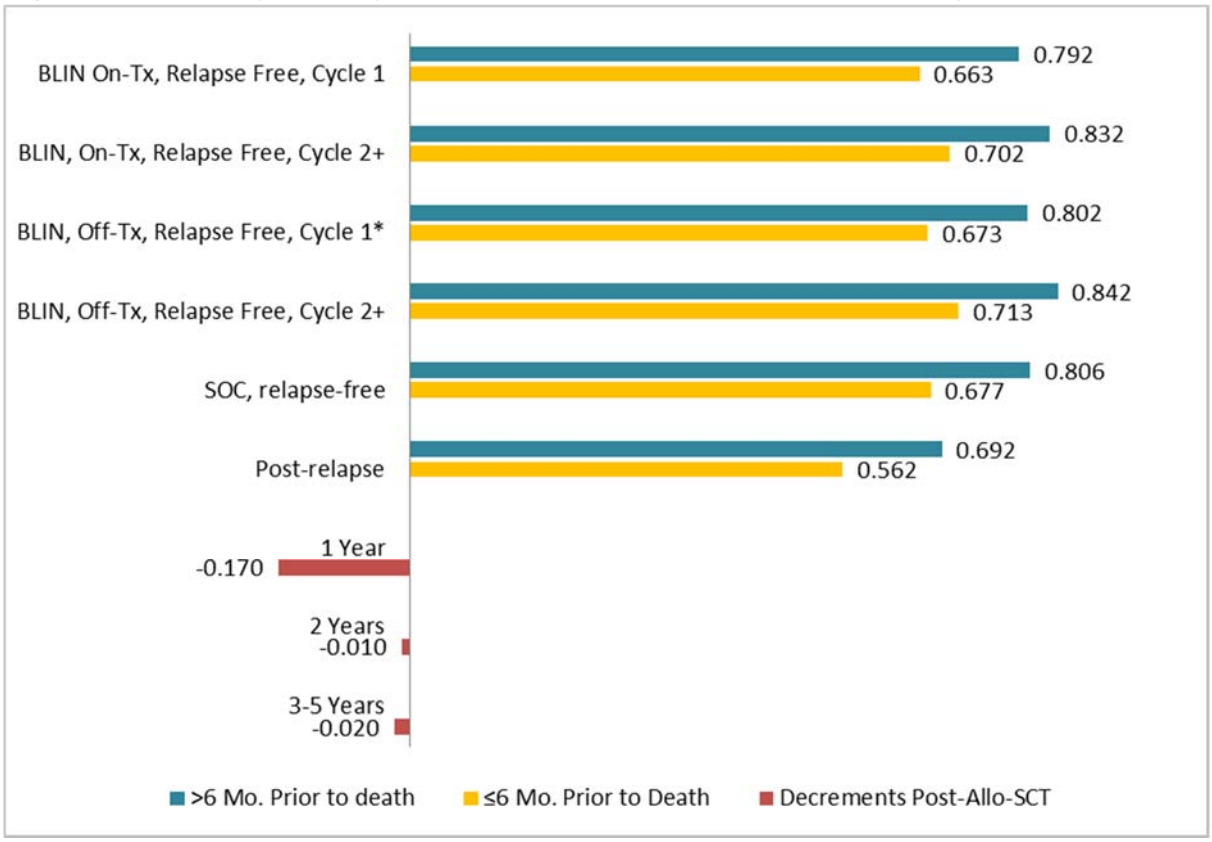
Table 47. Summary of utility values for cost-effectiveness analysis

State	Utility value: mean (standard error)	95% confidence interval	Reference in submission	Justification
Blinatumomab on-treatment relapse-free > 6 months prior to death				
Cycle 1	0.792	NA	B.3.4.1	Derived from regression equation, mean baseline utility value in BLAST and % MRD response
Cycle 2	0.832	NA	B.3.4.1	Derived from regression equation, mean baseline utility value in BLAST and % MRD response
Blinatumomab off-treatment relapse-free > 6 months prior to death				
Cycle 1*	0.802	NA	B.3.4.1	Derived from regression equation, mean baseline utility value in BLAST and % MRD response
Cycle 2	0.842	NA	B.3.4.1	Derived from regression equation, mean baseline utility value in BLAST and % MRD response
SoC relapse-free > 6 months prior to death	0.806	NA	B.3.4.1	Derived from regression equation, mean baseline utility value in BLAST and % MRD response
Post-relapse > 6 months prior to death	0.692	(0.649, 0.734)	B.3.4.1	From TOWER
Decrement in utility for <=6 months prior to death	-0.129	-0.1970, -0.0613	B.3.4.1	Derived from regression equation, mean baseline utility value in BLAST
Patients who survive for five years	Age and sex-matched	NA	B.3.4.1	Based on clinical expert opinion

	norms adjusted for long term decrement			that patients who survive for five years are likely to be cured of ALL, but may have residual decrements in HRQoL due to prior treatments
Decrement in utility value post-HSCT	Year 1: 0.17 Year 2: 0.01 Years 3-5: 0.02	Year 1: -0.22,0.56 Year 2: -0.16,0.18 Year 3-5: -0.11,0.15	B.3.4.1	Based on the differences in the mean utility value at these time points vs. at >5 years post HSCT (0.76) reported by Kurosawa et al. (2016).

Footnotes: *In the base-case, 100% of patients in blinatumomab start cycle 1; this utility value is not used in the base-case analysis.
Abbreviations: MRD: minimal residual disease; SoC, standard of care.

Figure 39. Summary of utility values used in the cost-effectiveness analysis



Abbreviations: BLIN: blinatumomab; Tx: treatment; SoC: standard of care; Mo: month; Allo-SCT: allogeneic haematopoietic stem cell transplantation.

B.3.5 *Cost and healthcare resource use identification, measurement and valuation*

In BLAST, blinatumomab was administered as a continuous IV infusion of 15 µg/m^2 per day over 4 weeks, followed by a 2-week treatment-free period, for up to four consecutive cycles. For the model, blinatumomab was assumed to be dosed at 28 µg/day, consistent with the dosing instructions in the Blincyto® label. The list price of blinatumomab was estimated to be £2,017 per 35 µg vial. Each 35 µg vial contains 28 µg of useable medication, thus patients were assumed to receive one vial of blinatumomab per day of treatment. The percentage of patients starting and completing each cycle was based on data for 73 patients in BLAST who were included in the PAS of the propensity-matched comparison of BLAST and the historical control (Table 48). Patients who discontinued treatment within a cycle were assumed to receive half the cost of a cycle, resulting in an average of 1.86 cycles of blinatumomab received per treatment.[65]

Table 48. Estimated percentage of patients starting and completing each cycle of blinatumomab

Cycle	Patients starting cycle (%)	Patients completing cycle (%)
1	100.00%	72.60%
2	65.75%	53.42%
3	31.51%	23.29%
4	17.81%	8.22%

It was assumed that blinatumomab would be administered on an inpatient basis for 4 days during the first cycle and the first 2 days of the second cycle, consistent with UK clinical expert opinion as described in NICE guidance TA450 for blinatumomab in previously treated Ph- B-precursor ALL.[75] The daily cost of hospitalisation for administration of blinatumomab was estimated to be £685.86, based on the ratio of the weighted average costs and weighted average length of stay for elective inpatient stays for the following HRG codes from the 2015-2016 National Schedule of Reference Costs:[101]

- SA24G – Acute Lymphoblastic Leukaemia with CC Score 5+
- SA24H – Acute Lymphoblastic Leukaemia with CC Score 2-4
- SA24J – Acute Lymphoblastic Leukaemia with CC Score 0-1

Patients receiving blinatumomab were assumed to receive daily home infusions for the remaining on-treatment days while not hospitalised. The *per diem* cost of the home infusion pump was estimated to be £3.84 which included the prorated cost of the pump, maintenance costs, and consumables (Table 49). Patients receiving blinatumomab were assumed to require an outpatient visit every four days to refill the pump at an outpatient infusion centre at a cost of £211.99, based on the NHS reference cost for HRG SB15Z: delivery of subsequent elements of a chemotherapy cycle.[101]

Table 49. Calculation of home infusion pump costs

	Cost (£)		
	Total	Per day	Per 28 days
Pump cost (5 years lifespan)	1,795	0.98	27.54

	Cost (£)		
	Total	Per day	Per 28 days
Annual maintenance costs	90	0.25	6.90
Consumables, pack of 20 (one every 4 days)	209	2.61	73.15
Total	2,094	3.84	107.59

Patients receiving blinatumomab were assumed to receive CSF prophylaxis with methotrexate, cytosine arabinoside (cytarabine), and dexamethasone every three months for up to two years or until death, relapse, or receipt of allogeneic HSCT, whichever occurs first (Table 50). Because the model does not keep track of the number of patients who have received HSCT and who subsequently relapsed, the percent of patients remaining alive and relapse-free and without HSCT at any point in time was approximated by the difference between RFS and the cumulative probability of pre-relapse HSCT. To the extent that some patients who received HSCT prior to relapse subsequently experienced relapse after HSCT, this assumption may lead to an underestimate of the time at risk for receiving CSF prophylaxis and hence the costs of such prophylaxis.

The unit costs of methotrexate, cytarabine, and dexamethasone were taken from eMit.[102] The cost of intrathecal administration of methotrexate was estimated to be £265.02 per 13-week cycle, based on the NHS Reference Cost for HRG code SB13Z: deliver more complex parenteral chemotherapy at first attendance.[101]

Table 50. Dosages and costs for CSF prophylaxis for patients receiving blinatumomab

Drug	Methotrexate	Cytarabine	Dexamethasone
Administration	Intrathecal	Intrathecal	Intrathecal
Dose per day of treatment	15 mg	40 mg	4 mg
Number of days administered per Cycle	1 / 13 weeks	1 / 13 weeks	1 / 13 weeks
Costs			
Cost per pack (£)	6.63	6.60	2.42
Units per pack	1	1	10
Mg per unit	1,000.0	2,000.0	3.3
Medication cost per unit (£)	6.63	6.60	0.24
Medication cost per mg (£)	0.01	0.00	0.07
Medication cost per day of administration (£)	0.10	0.13	0.24
Administration costs, per cycle*	265.02	–	–

Footnotes: Costs of administering cytarabine and dexamethasone assumed to be included in cost of administering methotrexate.

B.3.5.1 Intervention and comparators' costs and resource use

Patients receiving SoC were assumed to receive maintenance chemotherapy for up to two years or until relapse or allogeneic HSCT, whichever occurred first. As with the calculation of the costs of CSF prophylaxis for patients receiving blinatumomab, the percent of patients remaining alive

and relapse-free and without HSCT at any point in time was approximated by the difference between RFS and the cumulative probability of pre-relapse HSCT. To the extent that some patients who received HSCT prior to relapse subsequently experienced relapse after HSCT, this assumption may lead to an underestimate of the time at risk for receiving maintenance therapy and hence the costs of such therapy.

Maintenance chemotherapy was assumed to be comprised of 1.4 mg/m^2 of vincristine administered by IV infusion once every 13 weeks, 60 mg/m^2 of prednisolone taken orally five times per 13-week cycle, 75 mg/m^2 of mercaptopurine taken orally daily, 20 mg/m^2 of methotrexate taken orally weekly, and 12.5 mg of intrathecal methotrexate once every 13 weeks.[45] The unit costs of vincristine, prednisolone, and methotrexate (oral and IT) were from eMit.[102] The cost of mercaptopurine was not available from eMIT and was obtained from the BNF.[103] The cost of IV administration of vincristine was estimated to be £304.30 per 13-week cycle, based on the NHS Reference Cost for HRG code SB14Z: delivery of complex chemotherapy.[101] The cost of IT administration of methotrexate was estimated to be £265.02 per 13-week cycle, based on the NHS Reference Cost for HRG code SB13Z: deliver more complex parenteral chemotherapy at first attendance.[101] Administration costs for oral medications were assumed to be zero. The calculations of the dosage and costs for SoC maintenance are summarised in Table 51.

Table 51. Dosages and costs for SoC maintenance therapy

	Vincristine (IV)	Prednisolone (Oral)	Mercaptopurine (Oral)	Methotrexate (Oral)	Methotrexate Intrathecal
Dose per day of treatment	1.4 mg/m^2	60 mg/m^2	75 mg/m^2	20 mg/m^2	12.5 mg
Regimen	1 / 13 weeks	5 / 13 weeks	Daily	Weekly	1 x / 13 weeks
Cost per pack (£)	29.26	0.41	49.15	4.39	6.63
Units per pack	5	28	25	100	1
Mg per unit	2.0	5.0	50.0	2.5	1,000.0
Cost per unit (£)	5.85	0.01	1.97	0.04	6.63
Cost per mg (£)	2.93	0.00	0.04	0.02	0.01
Cost per day (medication only) (£)	7.76	0.33	5.59	0.67	0.08
Administration costs per cycle (£)	304.30	0	0	0	265.02

Footnotes: Cost calculations are based on a mean BSA of 1.89 m^2 calculated for the blinatumomab patients in the propensity matched analysis.
Abbreviations: SoC: standard of care; IV: intravenous.

B.3.5.2 Health-state unit costs and resource use

As noted above, the model considers the costs of the blinatumomab and SoC maintenance therapy, HSCT, salvage therapy, other ALL-related inpatient and outpatient care, and ALL-related terminal care costs. These costs are described in the sections below.

B.3.5.3 Adverse reaction unit costs and resource use

The costs of AEs were not considered explicitly in the model but were assumed to be captured in the costs of inpatient and outpatient care for the administration of blinatumomab and SoC maintenance therapy.

Therefore, the economic analysis presented in this submission does not include any additional AE unit costs or resource use.

B.3.5.4 Miscellaneous unit costs and resource use

Allogeneic Haematopoietic Stem Cell Transplant

Based on patients in the PAS of the propensity matched analysis of BLAST and the historical control (N=73 and N=182, respectively), 72.6% of blinatumomab patients from BLAST and 38.4% of SoC patients from the historical control study received allogeneic HSCT prior to relapse. Six-month probabilities of receiving HSCT prior to relapse were estimated to be 14.15% for blinatumomab 12.45% for SoC by calibrating the model to yield cumulative probabilities at 48 months (the approximate time of the last HSCT in BLAST or the historical control study) equal to 72.6% and 38.4%, respectively.

Data on post-relapse HSCT was not available from the BLAST and historical control studies. In the model, the probability of post-relapse HSCT was assumed to depend on receipt of HSCT prior to relapse and was estimated using data on receipt of HSCT by age and receipt of prior HSCT from the no prior salvage subgroup of a historical comparator study of patients with Ph-R/R B-precursor ALL (Protocol 20120310; Clinicaltrials.gov: NCT02003612).[104] For patients with and without prior HSCT, estimates of probability of receipt of allogeneic HSCT by age were weighted by the proportion of relapsing patients in BLAST within these age groups. Based on these data, it was estimated that 15.8% of patients with pre-relapse HSCT and 20.1% of those without pre-relapse HSCT will receive HSCT after relapse.

As noted above, the model does not keep track of the proportion of patients who receive HSCT and subsequently relapse. Accordingly, it is not possible to determine precisely the proportion of patients who relapse who have received prior HSCT. As an approximation, it was assumed that patients with pre-relapse HSCT would not relapse until all patients without pre-relapse HSCT have relapsed. Under this assumption, all relapses occurring prior to the point at which the RFS and cumulative HSCT curves cross are assumed to be among patients with no prior HSCT while all those occurring after that point are assumed to be amongst those with prior relapse. For discounting purposes, the cost of post-relapse HSCT was assumed to occur at the time of relapse. This assumption may lead to a slight overestimation of the discounted costs of post-relapse transplant.

The cost of HSCT was estimated based on an analysis conducted by the NHS Blood and Transplant Service.[105] This study included costs for initial treatment, and well as costs and probabilities of receipt of follow-up treatment for months 1–6, 7–12, 13–24, and > 24 months

post-transplant. Patients surviving > 24 months were assumed to receive daily cyclosporine. Costs from this study were adjusted to 2015/2016 prices using the PSSRU inflation index.[106] The cost of daily cyclosporine was estimated from the BNF.[103] Inputs used in the calculations of the costs of HSCT are shown Table 52.

Table 52. Inputs used in the calculation of the cost per patient receiving allogeneic HSCT

	Value
Initial treatment cost (£)	62,629
Follow-up treatment, percent of patients receiving (%)	
1-6 months	90
7-12 months	48
13-24 months	31
>24 months	20
Cost	
1-6 months cost (£)	30,186
7-12 months (£)	20,736
13-24 months (£)	14,963
> 24 months, cyclosporine	
Mg per day	100
Cost per tab (£)	0.85
Mg per tab	50.00

Abbreviations: HSCT: haematopoietic stem cell transplant

Salvage Therapy

In the base-case, it was assumed that all patients would receive SoC salvage therapy upon relapse. The cost of SoC salvage therapy was estimated using an economic model used in the manufacturer's submission in response to STA of blinatumomab for previously treated B-precursor acute lymphoblastic leukaemia (STA 1804).[107] This model was used to calculate the costs of first and second salvage therapy for the subgroup of patients with no prior salvage therapy assuming that all patients who relapse would receive first-line salvage therapy, that 37.0% of patients who relapse after first-line salvage therapy would receive second-line salvage therapy, and that the cost per course of salvage therapy is £16,175, based on medication and administration costs for FLAG-IDA. Based on these assumptions, the model generated estimate of the cost of first and second-line salvage therapy (discounted to time of initiation of first-line salvage) is £21,905. As noted above, ALL-related costs after 60 months, including the cost of salvage therapy, were assumed to be zero in the base-case.

As noted above, it was assumed in the base-case that salvage therapy for patients experiencing relapse would be SoC chemotherapy, consistent with the salvage therapy received among patients in the BLAST trial and the historical control study. While the model estimates of OS are internally consistent with this assumption, it does not reflect the likely use of blinatumomab as salvage therapy given the recent NICE guidance.[75] Because modelling of a "counterfactual" scenario in which relapsing patients would receive blinatumomab was not feasible using the data from BLAST and the historical control study within the PartSA model structure, a scenario analysis was conducted in which incremental costs and QALYs generated by the model were adjusted to reflect the difference between treatments in (1) the percent of patients receiving

salvage therapy and (2) the incremental costs and QALYs associated with blinatumomab versus SoC chemotherapy salvage. The latter was estimated based on the economic model used in the evaluation of blinatumomab versus SoC salvage therapy in R/R ALL, focusing on the subgroup of patients without prior salvage therapy.[75] This scenario is described in greater detail in Section B.3.8 below.

Other Inpatient and Outpatient Costs

SoC patients, and blinatumomab patients who are no longer receiving blinatumomab treatment, were assumed to incur additional inpatient and outpatient costs that were assumed be dependent on MRD response. In the base-case, inpatient and outpatient healthcare resource utilisation (HRU) by MRD response was based on results of face-to-face interviews of two UK experts – this approach was considered appropriate given the rare and complex nature of this disease area. Nevertheless, a follow-up, larger multinational online survey that was also conducted to gather more information on patterns of testing for MRD response.

The results for the 20 UK physicians participating in the online survey were used only in a scenario analysis (see Section B.3.8). While the sample was larger for the online survey, the results from the face-to-face interviews are likely to be more robust, as the face-to-face discussion ensured that the clinicians understood well the questions regarding resource utilisation. In fact, the distributions of results from the online survey for questions regarding HRU suggested that many physicians participating in the online survey did not adequately understand the questions and that their responses were, therefore, potentially biased. Mean inpatient days and physician visits per month from the face-to-face interviews and online survey of UK physicians are reported in Table 53.

Table 53. Mean inpatient and outpatient HRU per month by MRD response from face-to-face interviews and online survey of UK physicians

	Face-to-Face Interview (N=2)		Online Survey (N=20)	
Services	**MRD +**	**MRD-**	**MRD +**	**MRD-**
Inpatient days	1.75	0.06	3.10	2.33
Outpatient				
Haematologist	2.000	1.500	1.167	0.917
Radiologist	0.417	0.250	0.250	0.083
Other specialist	0.500	0.250	0.250	0.250
General physician	0.750	0.417	0.833	0.500

Abbreviations: HRU: healthcare resource utilisation; MRD: minimal residual disease; UK: United Kingdom.

The probability of MRD response for blinatumomab patients was estimated to be 83.6% based on the rate of MRD response after cycle 2 among patients in BLAST who were included in the PAS of the propensity matched analysis of BLAST and the historical control study. As described above, the probability of MRD response for patients receiving SoC maintenance therapy was assumed to be 8%. Because data on HRU after relapse was not available from the survey or any other comparable source, it was assumed that HRU post-relapse would be independent of initial treatment and the same as pre-relapse HRU for patients in haematological remission but with no MRD response. As with other ALL-related costs, in the base-case, these costs were assumed to be incurred only over the first five years of the model.

The cost per inpatient day was assumed to be the same as that for blinatumomab administration (£685.86 per day), which was based on NHS Reference Costs (as described above). The costs of the visit to haematologists, radiologists, and other specialists were estimated to be £166.03, £51.35, and £162.84, respectively, based on NHS Reference Costs for non-admitted face-to-face attendances for consultations for clinical haematology (HRG code WF01A-303), diagnostic imaging (HRG code WF01A-812) and medical oncology (HRG code WF01A-370).[101] The cost of a visit to a general physician was estimated to be £36.00 based on the general practitioner unit cost from the PSSRU.[106] Unit costs for inpatient days and visits are summarized in Table 54.

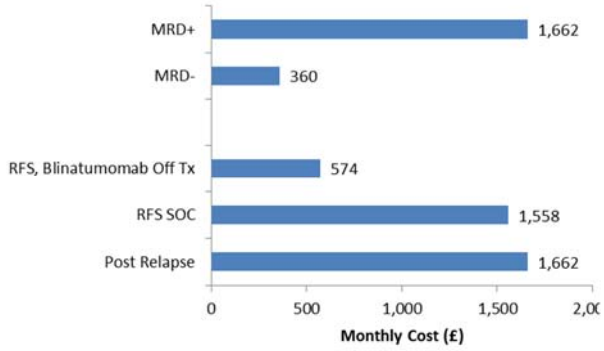
Table 54. Other inpatient and outpatient costs

Other costs	Cost (£)
Inpatient day	685.86
Outpatient, per visit	
Haematologist	166.03
Radiologist	51.35
Other specialist	162.84
General Physician	36.00

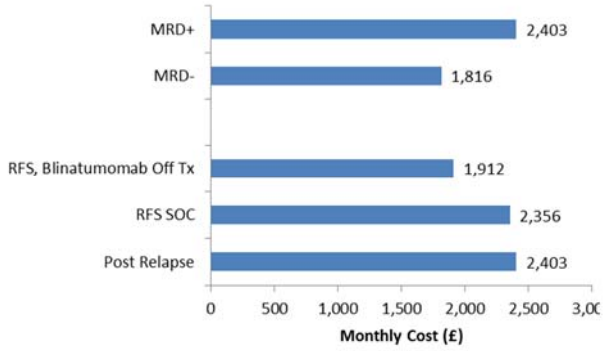
Monthly costs by MRD response, treatment, and relapse using HRU from the face-to-face interviews (base-case) and online survey (sensitivity analyses) are shown in Figure 40.

Figure 40. Mean monthly other ALL-related inpatient and outpatient costs by MRD response, treatment, and relapse using HRU from face-to-face interviews and online survey

A. HRU from Face-to-Face Interviews



B. HRU from Online Survey



Abbreviations: ALL: acute lymphoblastic leukaemia; HRU: healthcare resource utilisation; MRD: minimal residual disease; RFS: relapse-free survival; Tx: treatment; SoC: standard of care.

Terminal Care

ALL-related terminal care costs were estimated to be £8,602 based on the average length of stay in the hospital for terminally ill patients (8 weeks) as reported in a recent report by the King`s Fund,[108] and the average cost of end-of-life care (£145 per day) from Marie Curie,[109] adjusted to 2015/2016 values using the pay and prices index for Hospital and Community Health Service (HCHS) from the PSSRU [106]. As noted above, ALL-related costs after 60 months, including the cost of terminal care, were assumed to be zero in the base-case.

B.3.6 *Summary of base-case analysis inputs and assumptions*

B.3.6.1 Summary of base-case analysis inputs

The inputs used in the base-case analysis are reported in Table 55.

Table 55. Summary of variables applied in the economic model

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
Analytic variables			
Timeframe (years)	50	NA	B.3.2.2
Annual discount rate for costs	3.5%	NA	
Annual discount rate for effectiveness	3.5%	NA	
ICER threshold	50,000	NA	
Patient characteristics			
Starting age (years)	45.4	45.4 to 45.4	B.2.3.3
Percent male	56%	56% to 56%	
Mean BSA	1.9	1.9 to 1.9	
Efficacy			
MRD response rate			
Blinatumomab	83.6%	74.2% to 91.2% (Beta)	B.2.6
SoC	8.0%	NA	
RFS distribution	Unrestricted Gompertz	NA	
OS distribution	Lognormal Mixture cure	NA	
Proportion of RFS events that are deaths			
Blinatumomab	47.1%	NA	B.2.6
SoC	8.5%	NA	
RR of death versus gen. pop. mort	4	NA	B.3.2.2
Mos. after which gen. population utility values are used	60	NA	
Mos. after which pre-relapse other in/outpatient costs are set to zero	60	NA	

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
Mos. after which post-relapse other in/outpatient costs are set to zero	60	NA	
Mos. after which salvage therapy costs are set to zero	60	NA	
Mos. after which terminal care costs are set to zero	60	NA	
Mos. after which terminal decrement in utility is set to zero	60	NA	
Costs			
Blinatumomab			
Cost per vial (list price)	2,017.00	NA	B.3.5
PAS discount	[redacted]	NA	
Days per bag change	4	NA	
Inpatient costs			
Inpatient days per cycle received			
Cycle 1	4	NA	B.3.5
Cycle 2	2	NA	
Cycle 3	0	NA	
Cycle 4	0	NA	
Cost per inpatient day	685.86	410.66 to 1,078.08 (Lognormal)	
Outpatient costs			
Probability of receiving infusions in outpatient infusion centre			
Cycle 1	1	NA	B.3.5
Cycle 2	1	NA	
Cycle 3	1	NA	
Cycle 4	1	0 to 0	
Cost per visit to outpatient infusion centre	211.99	126.93 to 333.22 (Lognormal)	
Pump costs			

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
Total cost of pump per patient	7,008.00	4,196.10 to 11,015.72 (Lognormal)	B.3.5
Useful life expectancy of pump (years)	5	NA	
Duration of therapy			
% Starting cycle			
Cycle 1	100.0%	100.0% to 100.0% (Beta)	B.3.5
Cycle 2	65.8%	54.5% to 76.2% (Beta)	
Cycle 3	31.5%	21.4% to 42.6% (Beta)	
Cycle 4	17.8%	9.9% to 27.4% (Beta)	
% Completing cycle			
Cycle 1	72.6%	61.8% to 82.2% (Beta)	B.3.5
Cycle 2	53.4%	41.9% to 64.7% (Beta)	
Cycle 3	23.3%	14.3% to 33.6% (Beta)	
Cycle 4	8.2%	3.1% to 15.5% (Beta)	
CSF prophylaxis			
Dexamethasone (intrathecal)			
Dose per day of treatment	4	NA	B.3.5
Unit of measurement	mg	NA	
No. of days administered per cycle	1	NA	
Cycle length (weeks)	13.0	NA	
Duration of treatment (weeks)	104.3	NA	
Cost			
Cost per pack (£)	3.21	NA	B.3.5
Units per pack	5	NA	
Mg per unit	6.6	NA	

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
Administration costs, per cycle	265.02	158.68 to 416.58 (Lognormal)	
Methotrexate (intrathecal)			
Dose per day of treatment	15	NA	B.3.5
Unit of measurement	mg	NA	
No. of days administered per cycle	1	NA	
Cycle length (weeks)	13.0	NA	
Duration of treatment (weeks)	104.3	NA	
Cost			
Cost per pack (£)	9.32	NA	B.3.5
Units per pack	1	NA	
Mg per unit	1,000.0	NA	
Administration costs, per cycle	0.00	NA	
Cytosine arabinoside (intrathecal)			
Dose per day of treatment	40	NA	B.3.5
Unit of measurement	mg	NA	
No. of days administered per cycle	1	NA	
Cycle length (weeks)	13.0	NA	
Duration of treatment (weeks)	104.3	NA	
Cost			
Cost per pack (£)	8.17	NA	B.3.5
Units per pack	1	NA	
Mg per unit	2,000.0	NA	
Administration costs, per cycle	0.00	NA	
SoC			

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
Vincristine (IV)			
Dose per day of treatment	1	NA	B.3.5.1
Unit of measurement	mg/m^2	NA	
No. of days administered per cycle	1	NA	
Cycle length (weeks)	13.0	NA	
Duration of treatment (weeks)	104.3	NA	
Cost			
Cost per pack (£)	15.71	NA	B.3.5.1
Units per pack	5	NA	
Mg per unit	1.0	NA	
Administration costs, per cycle	304.30	182.20 to 478.32 (Lognormal)	
Prednisolone (oral)			
Dose per day of treatment	60	NA	B.3.5.1
Unit of measurement	mg/m^2	NA	
No. of days administered per cycle	5	NA	
Cycle length (weeks)	13.0	NA	
Duration of treatment (weeks)	104.3	NA	
Cost			
Cost per pack (£)	0.42	NA	B.3.5.1
Units per pack	28	NA	
Mg per unit	5.0	NA	
Administration costs, per Cycle	0.00	NA	
Mercaptopurine (oral)			
Dose per day of treatment	75	NA	B.3.5.1

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
Unit of measurement	mg/m^2	NA	
No. of days administered per cycle	28	NA	
Cycle length (weeks)	4.0	NA	
Duration of treatment (weeks)	104.3	NA	
Cost			
Cost per pack (£)	49.15	NA	B.3.5.1
Units per pack	25	NA	
Mg per unit	50.0	NA	
Administration costs, per Cycle	0.00	NA	
Methotrexate (oral)			
Dose per day of treatment	20	NA	B.3.5.1
Unit of measurement	mg/m^2	NA	
No. of days administered per cycle	4	NA	
Cycle length (weeks)	4.0	NA	
Duration of treatment (weeks)	104.3	NA	
Cost			
Cost per pack (£)	4.63	NA	B.3.5.1
Units per pack	100	NA	
Mg per unit	2.5	NA	
Administration costs, per Cycle	0.00	NA	
Methotrexate (intrathecal)			
Dose per day of treatment	13	NA	B.3.5.1
Unit of measurement	mg	NA	
No. of days administered per cycle	1	NA	

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
Cycle length (weeks)	13.0	NA	
Duration of treatment (weeks)	104.3	NA	
Cost			
Cost per pack (£)	9.32	NA	B.3.5.1
Units per pack	1	NA	
Mg per unit	1,000.0	NA	
Administration costs, per cycle	265.02	158.68 to 416.58 (Lognormal)	
Probability of receiving HSCT blinatumomab pre-relapse			
Month			
1-6	14.1%	7.4% to 22.6% (Beta)	B.3.5.4
7-12	14.1%	7.4% to 22.6% (Beta)	
13-18	14.1%	7.4% to 22.6% (Beta)	
19-24	14.1%	7.4% to 22.6% (Beta)	
25-30	14.1%	7.4% to 22.6% (Beta)	
31-36	14.1%	7.4% to 22.6% (Beta)	
37-42	14.1%	7.4% to 22.6% (Beta)	
43-48	14.1%	7.4% to 22.6% (Beta)	
49+	0.0%	NA (Beta)	
Probability of receiving HSCT, SoC pre-relapse			
Month			
1-6	12.5%	9.4% to 15.8% (Beta)	B.3.5.4
7-12	12.5%	9.4% to 15.8% (Beta)	
13-18	12.5%	9.4% to 15.8% (Beta)	
19-24	12.5%	9.4% to 15.8% (Beta)	

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
25-30	12.5%	9.4% to 15.8% (Beta)	
31-36	12.5%	9.4% to 15.8% (Beta)	
37-42	12.5%	9.4% to 15.8% (Beta)	
43-48	12.5%	9.4% to 15.8% (Beta)	
49+	0.0%	NA (Beta)	
Probability of receiving HSCT, post-relapse			
With no prior HSCT	20.1%	13.8% to 27.1% (Beta)	B.3.5.4
With prior HSCT	15.8%	10.2% to 22.3% (Beta)	
HSCT cost			
Initial treatment	62,629.00	37,499.68 to 98,445.17 (Lognormal)	B.3.5.4
Follow-up			
Percent of patients receiving			
0-6 Mos.	90.0%	NA	B.3.5.4
7-12 Mos.	48.0%	NA	
13-24 Mos.	31.0%	NA	
>24 Mos., cyclosporine	20.0%	NA	
Cost 1-6 Mos. (£)	30,186.00	18,074.14 to 47,448.72 (Lognormal)	
Cost 7-12 Mos. (£)	20,736.00	12,415.87 to 32,594.47 (Lognormal)	
Cost 13-24 Mos. (£)	14,963.00	8,959.23 to 23,520.02 (Lognormal)	
Cost after 24 Mos., cyclosporine (£)			
Days of use per cycle	30.40	NA	B.3.5.4
Days per cycle	30.40	NA	
Salvage chemotherapy			
Proportion receiving each Treatment			

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
Blinatumomab patients			
Multi-agent chemotherapy	100.0%	NA	B.3.5.4
Blinatumomab	0.0%	NA	
SoC patients			
Multi-agent chemotherapy	100.0%	NA	B.3.5.4
Blinatumomab	0.0%	NA	
Cost per course			
Multi-agent chemotherapy	21,905.39	13,116.05 to 34,432.61 (Lognormal)	B.3.5.4
Blinatumomab	97,176.00	58,185.01 to 152,748.86 (Lognormal)	
Other inpatient costs			
No. of inpatient days per Month			
MRD+	1.75	NA	B.3.5.4
MRD-	0.06	NA	
Cost per inpatient day	685.86	410.66 to 1,078.08 (Lognormal)	
Other outpatient costs			
No. of visits to haematologist per month			
MRD+	2.00	NA	B.3.5.4
MRD-	1.50	NA	
Cost per visit to haematologist	166.03	99.41 to 260.98 (Lognormal)	
No. of visits to radiologist per month			
MRD+	0.42	NA	B.3.5.4
MRD-	0.25	NA	
Cost per visit to radiologist	51.35	30.75 to 80.71 (Lognormal)	
No. of visits to general physician per month			

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
MRD+	0.75	NA	B.3.5.4
MRD-	0.42	NA	
Cost per visit to general physician	36.00	21.56 to 56.59 (Lognormal)	
No. of visits to other specialist per month			
MRD+	0.50	NA	B.3.5.4
MRD-	0.25	NA	
Cost per visit to general physician	162.84	97.50 to 255.97 (Lognormal)	
Terminal care (for patients not cured)	8,833.84	5,289.34 to 13,885.73 (Lognormal)	
Utility Inputs			
TOWER EQ-5D analysis			
Post-relapse, mean	0.692	0.693 to 0.690 (Lognormal (Utility))	B.3.4.1
BLAST EQ-5D analysis			
Baseline, mean	0.809	0.817 to 0.800 (Lognormal (Utility))	B.3.4.1
GLM/GEE regression analysis covariates			
Intercept	0.353	NA (Var-Covar Matrix)	B.3.4.1
Baseline	0.543	NA (Var-Covar Matrix)	
Off-Treatment Relapse-Free	0.010	NA (Var-Covar Matrix)	
MRD Response	0.047	NA (Var-Covar Matrix)	
Terminal Decrement	-0.129	NA (Var-Covar Matrix)	
Gen. pop.			
Male			
18 – 24	0.940	NA	B.3.4.1
25 – 34	0.930	NA	
35 – 44	0.910	NA	

Variable	Value	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
45 – 54	0.840	NA	
55 – 64	0.780	NA	
65 – 74	0.780	NA	
75 – 100	0.750	NA	
Female			
18 – 24	0.940	NA	B.3.4.1
25 – 34	0.930	NA	
35 – 44	0.910	NA	
45 – 54	0.850	NA	
55 – 64	0.810	NA	
65 – 74	0.780	NA	
75 – 100	0.710	NA	
Long-term decrement versus general population	0.02		

Abbreviations: CI: confidence interval; ICER: incremental cost effectiveness ratio; BSA: body surface area; MRD: minimal residual disease; SoC: standard of care; RFS: relapse-free survival; OS: overall survival; RR: relative risk: Mos: months; CSF: cerebrospinal fluid; IV: intravenous: HSCT: haematopoietic stem cell transplantation; EQ-5D: EuroQol 5 dimensions questionnaire; GLM: generalised linear model; GEE: generalised estimating equations.

B.3.6.2 Assumptions

Key modelling assumptions and their justifications are listed below (Table 56).

Table 56. Key Modelling Assumptions

Model input and cross reference	Source/assumption	Justification
The population of interest is patients with Ph- B-precursor ALL in MRD+ CR1 (Section B.3.6.2, page 144)	Data limitation arising from the historical comparator study	Blinatumomab will be used as early as possible in the treatment pathway, i.e. in CR1 patients, where the benefits of treatment are likely to be greatest. Although the cost effectiveness evidence does not consider the Ph+ population or later remission states, due to the substantial unmet need across all sub-populations blinatumomab should be considered for use in alignment with its full anticipated marketing authorisation.
The appropriate comparator of interest is SoC maintenance therapy (Section B.3.6, page 144)	Conservative assumption based on eligibility criteria of BLAST/historical cohort and clinical practice based on UKALL14 protocol	In BLAST, haematological CR was defined as < 5% blasts in bone marrow after at least 3 intense chemotherapy blocks. Accordingly, it is appropriate to assume that patients receiving blinatumomab will have already received induction, intensification, and consolidation therapy. A conservative assumption was therefore made to only include the cost of ongoing maintenance chemotherapy to avoid counting costs of treatment that may already have been received.
All patients who relapse before 60 months receive multi-agent chemotherapy (FLAG-IDA) as salvage therapy	Consistent with SoC salvage therapy received among patients in BLAST and historical control study.	Model estimates of OS are internally consistent with SoC salvage although do not reflect the likely use of blinatumomab as salvage therapy given the recent NICE TA450 guidance. A key scenario analysis was therefore conducted to assess the impact of salvage treatment with blinatumomab in the SoC arm to better reflect current clinical practice in the UK.
Mortality rates for blinatumomab and SoC will never be less than age- and sex-match general population mortality rates adjusted for the increase in long-term mortality due to exposure to radiotherapy, cytotoxic chemotherapy, and HSCT (Section B.3.6, page 144)	Conservative assumption, similar to NICE TA450[22]	Mortality rates will decline initially as patients who are not cured die, and then are expected to increase over time due to increasing non–disease-related mortality in cured patients.

Utility values after 5 years are the same for patients receiving blinatumomab and SoC and assumed to be equal to UK general population norms for EQ-5D adjusted for long-term decrement in HRQoL due to exposure to radiotherapy, cytotoxic chemotherapy, and HSCT (Section B.3.6, page 144)	Clinical expert opinion Adjusted from NICE TA450 to reflect the earlier position in the treatment pathway (5 years here, 4 years in TA450)[22]	Patients surviving for 5 years are likely to be cured of ALL and to no longer suffer from disease-related decrements in HRQoL. Consequently, utility values will be the same for patients receiving blinatumomab and FLAG-IDA. Over time, utility values will decrease due to age-related reductions in HRQoL. Clinical expert opinion supports the notion of cure at 5 years if not before.
ALL-related costs (excluding follow-up costs associated with HSCT conducted previously) are zero in both groups after 5 years (Section B.3.6, page 144)	Clinical expert opinion	This assumption is based on clinical expert opinion that patients who survive for five years are likely to be cured and will no longer require ALL treatment (excluding follow-up costs associated with HSCT conducted previously).
Costs of AEs are captured in costs of inpatient and outpatient administration of medications (Section B.3.6, pages 144–145)	Aligned with NICE TA450[22]	Since blinatumomab is administered initially in hospital, the treatment of AEs is likely to be provided during the hospital stay and therefore included in the hospitalisation cost. As patients are assumed to visit outpatient infusion centres every 4 days when receiving the drug out of hospital, it is likely AEs could be managed during these scheduled visits.
For patients in the RF state, utility values are dependent whether patients are on vs. off blinatumomab treatment, MRD response, and time from death (Section B.3.6, page 145)	BLAST	This assumption is supported by regression analyses on EQ-5D utility values among patients in BLAST. Although the coefficients for off vs. on treatment and MRD response were not statistically significant, they were directionally consistent with expectations (i.e., higher utility values for patients off treatment and with MRD response).
Patients receiving blinatumomab will be hospitalised for four days for the first cycle and 2 days for the second cycle. Remaining cycles will be administered on an outpatient basis (Section B.3.6, page 145)	Aligned with NICE TA450[22]	This assumption is consistent with the appraisal committee's preferred assumption in NICE TA450 recommending blinatumomab for previously treated Ph- ALL.

Abbreviations: Ph: Philadelphia chromosome; ALL: acute lymphoblastic leukaemia; CR1: first haematological complete remission; MRD: minimal residual disease; SoC: standard of care; HSCT: haematopoietic stem cell transplantation; UK: United Kingdom; EQ-5D: EuroQol 5 dimensions questionnaire; HRQoL: health-related quality of life; FLAG-IDA: fludarabine, cytarabine, granulocyte colony-stimulating factor, idarubicin; AE: adverse events; RF: relapse-free; TA: technology appraisal.

B.3.7 *Base-case results*

Base-case results are presented in the following sub-section. Clinical outcomes and dissagregated results are presented in Appendix J.

B.3.7.1 Base-case incremental cost-effectiveness analysis results

Base-case results for the cost effectiveness of blinatumomab versus SoC in adult patients with Ph- B-precursor ALL are reported in Table 57. Blinatumomab was projected to yield 3.58 more discounted life-years (LYs) and 2.95 more discounted QALYs than SoC. Total costs were estimated to be [redacted] higher with blinatumomab than with SoC. The ICER for blinatumomab versus SoC was therefore estimated to be £28,524 per QALY gained. As described in Section B.3.8 below, mean ICER from the PSA (calculated as the ratio of the mean incremental costs to the mean incremental QALYs) is £29,673 per QALY gained.

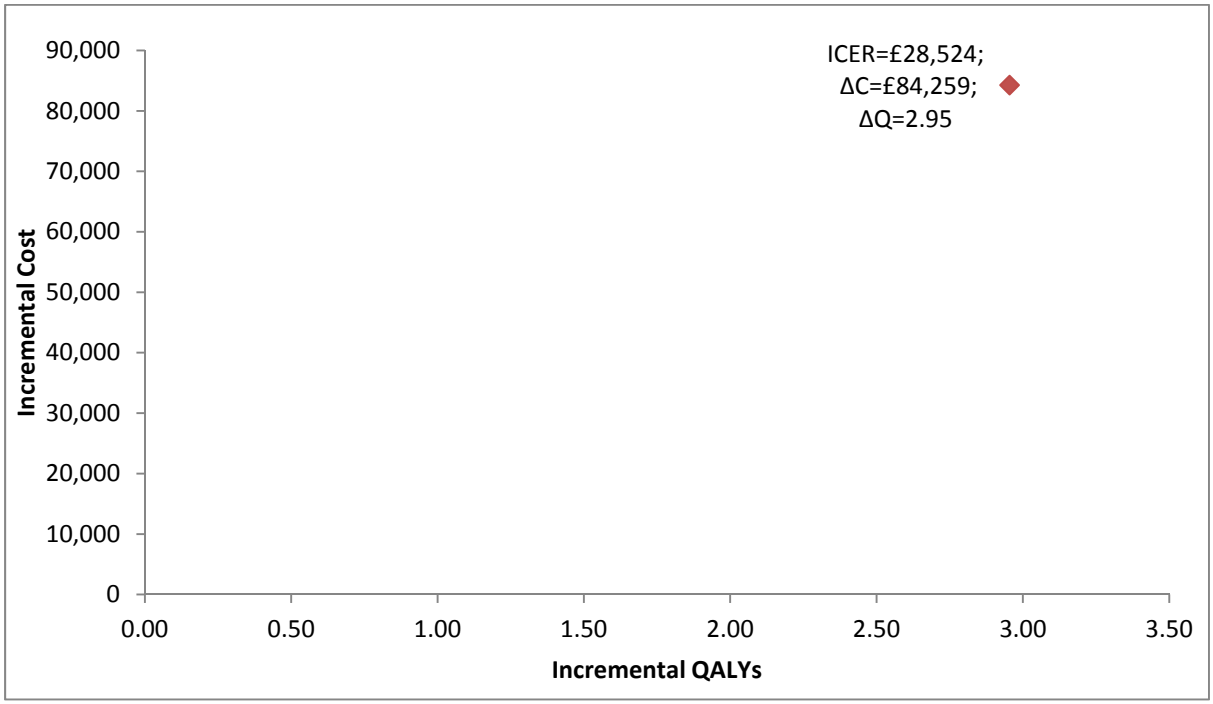
Table 57: Base-case results

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
SoC	[redacted]	5.51	4.14				
Blinatumomab	[redacted]	9.09	7.10	84,259	3.58	2.95	28,524

Abbreviations: ICER: incremental cost-effectiveness ratio; LYG: life years gained; QALYs: quality-adjusted life years; SoC: standard of care.

Incremental costs and QALYs with blinatumomab versus SoC are plotted on the cost-effectiveness plane in Figure 41. Also shown on the figure is the line representing a willingness-to-pay (WTP) threshold of £50,000 per QALY gained. The co-ordinates for the base-case estimate of the ICER is below the line suggesting that blinatumomab is a cost-effective use of healthcare resources given this threshold.

Figure 41. Incremental costs and QALYs with blinatumomab versus SoC



Abbreviations: SoC: standard of care; ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life-years; WTP: willingness to pay.

B.3.7.2 Key scenario analysis – blinatumomab as salvage treatment in SoC

Results for a key scenario analysis where patients who relapse on SoC receive blinatumomab are presented in Table 58. This scenario was conducted to more accurately reflect current UK clinical practice given the recent NICE TA450 guidance recommending the use of blinatumomab in patients with relapsed/refractory Ph- B-Cell ALL. In this scenario, it was assumed that 70% of patients who relapse while receiving SoC maintenance would receive blinatumomab as salvage therapy, based on unpublished forecasts of market share from Amgen.

This scenario was implemented using the incremental costs and QALYs from the company evidence submission in response to the recent NICE STA of blinatumomab with R/R B-precursor ALL.[107] Specifically, the costs of blinatumomab salvage [redacted]) was estimated by adding to the estimated cost of multi-agent chemotherapy salvage ([redacted]) the incremental costs of blinatumomab versus FLAG-IDA salvage for the sub-population of patients with no prior salvage therapy, using the PAS discount [redacted]). To account for the beneficial effects of blinatumomab salvage, the 70% of patients receiving blinatumomab salvage were assigned the estimated discounted incremental life-year gain of 2.40 and a QALY gain with blinatumomab versus FLAG-IDA salvage (1.98 QALYs gained). For discounting, the life year and QALY gains were assigned at the time of relapse. In this scenario, blinatumomab was projected to yield 2.31 more discounted life-years (LYs) and 1.91 more discounted QALYs than SoC. Total costs were

estimated to be [redacted] higher with blinatumomab than with SoC. The ICER for blinatumomab versus SoC was therefore estimated to be £17,420 per QALY gained.

Table 58. Key scenario results – blinatumomab as salvage tx for SoC – B.3.7.1 (page 146)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental. costs (£)	Incremental LYG	Incremental QALYs	Incremental ICER (£/QALY)
SoC	[redacted]	6.78	5.19				
Blinatumomab	[redacted]	**9.09**	**7.10**	**33,473**	**2.31**	**1.91**	**17,420**

Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years

B.3.8 *Sensitivity analyses*

B.3.8.1 Probabilistic sensitivity analysis

Probabilistic sensitivity analyses (PSAs) were generated based on 10,000 Monte-Carlo simulations with sampling from the distributions of parameter estimates for which distributional information was available. Parameters of survival distributions were sampled from bootstrap distributions derived from the source data (BLAST and historical control).

Results of PSAs for the comparison of blinatumomab versus SoC are summarised in Table 59.

Table 59. Results of PSA of blinatumomab versus SoC

Outcome	Blinatumomab	SoC	Incremental
Life years (not discounted)			
Mean	13.28	7.76	5.52
SD	2.31	1.28	2.44
Median	13.49	7.87	5.52
95% LCL	7.47	5.12	0.00
95% UCL	17.02	10.01	9.93
QALYs (discounted)			
Mean	6.96	4.11	2.85
SD	1.10	0.60	1.19
Median	7.04	4.15	2.86
95% LCL	4.31	2.87	0.21
95% UCL	8.78	5.19	5.04
Cost (discounted) (£)			
Mean	[redacted]	[redacted]	[redacted]
SD	[redacted]	[redacted]	[redacted]
Median	[redacted]	[redacted]	[redacted]
95% LCL	[redacted]	[redacted]	[redacted]
95% UCL	[redacted]	[redacted]	[redacted]

Abbreviations: PSA: probabilistic sensitivity analyses; SoC: standard of care; SD: standard deviation; LCL: lower confidence limit; UCL: upper confidence limit; QALY: quality-adjusted life year.

Box and whisker plots of the PSA results for undiscounted life-years, QALYs, and costs are shown in Figure 42.

Figure 42. Box and whisker plots for distributions of LY, QALYs and costs from PSA

Life Years **QALYs**

Costs

Abbreviations: LY: life years; QALY: quality-adjusted life year; PSA: probabilistic sensitivity analyses; SoC: standard of care.

The results of the PSA with respect to cost-effectiveness are summarised in Table 60. Given an ICER threshold of £50,000/QALY, the mean NMB was £57,855. The mean ICER from the PSA was £29,673.

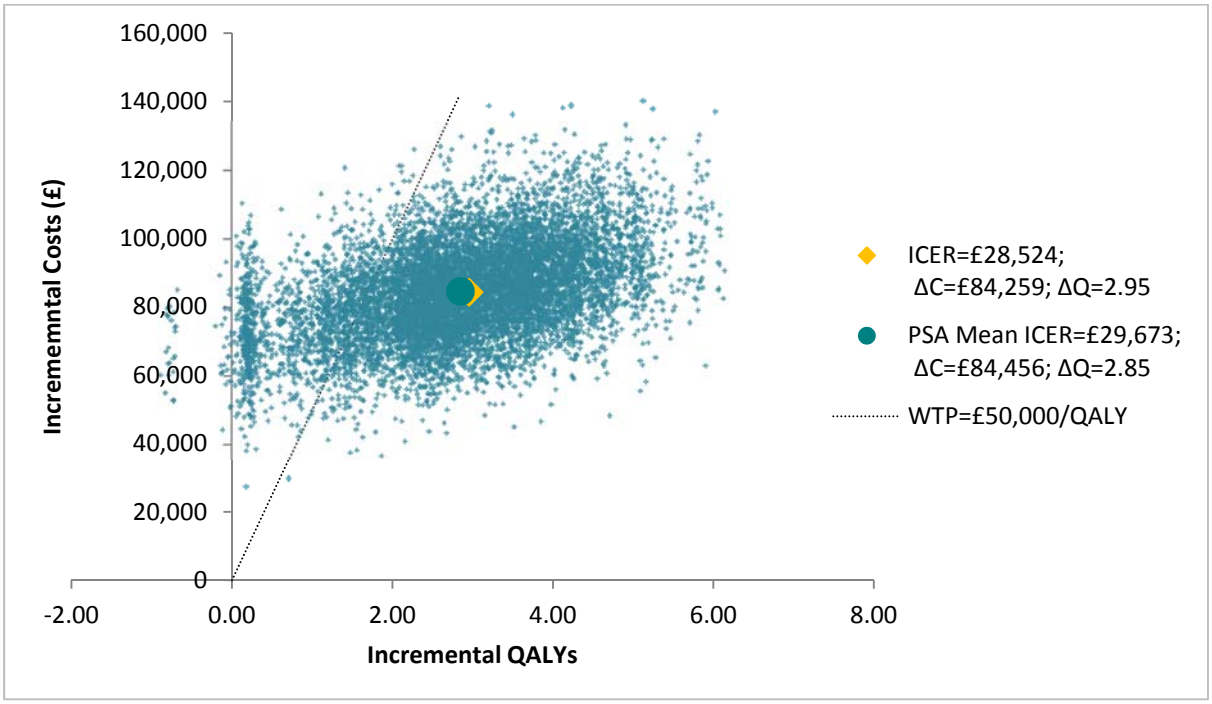
Table 60. Cost-effectiveness results from probabilistic sensitivity analyses

	Value
Percent of simulations in quadrant of CE plane	
Northeast (more costly and more effective)	99.5%
Southeast (dominant)	0.0%
Southwest (less costly and less effective)	0.0%
Northwest (dominated)	0.5%
NMB (WTP = £50,000 per QALY) (£)	
Mean	57,855
SD	54,845
Median	59,472
95% LCL	-61,944
95% UCL	158,790
Probability that therapy is preferred (WTP = £50,000)	
Blinatumomab	85.5%
SoC	14.5%
PSA mean ICER (ratio of mean incremental cost to mean incremental QALYs) (£)	29,673

Abbreviations: CE: cost-effectiveness; NMB: net monetary benefit; WTP: willingness to pay threshold; QALY: quality-adjusted life year; SD: standard deviation; LCL: lower confidence limit; UCL: upper confidence limit; SoC: standard of care; PSA: probabilistic sensitivity analyses; ICER: incremental cost-effectiveness ratio.

A scatter plot of the joint distribution of incremental costs and incremental QALYs from the PSA is shown in Figure 43. It should be noted that the correlation of the incremental costs and QALYs is relatively modest (Pearson correlation coefficient = 0.44). This reflects that the blinatumomab medication and administration costs are modelled independently of clinical outcomes.

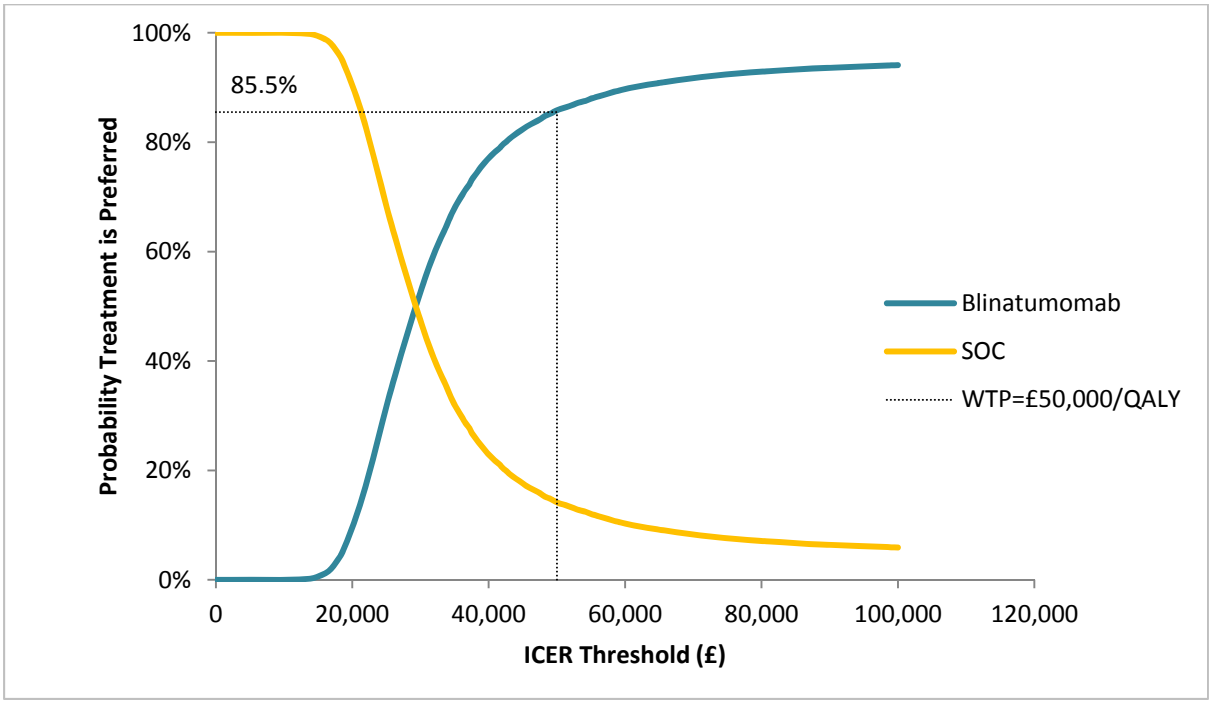
Figure 43. Scatter plot of simulations on cost-effectiveness plane



Abbreviations: ICER: incremental cost-effectiveness ratio; WTP: willingness to pay threshold; QALY: quality-adjusted life year.

Cost-effectiveness acceptability curves for blinatumomab and SoC care shown in Figure 44. The probability that blinatumomab is preferred was estimated to be 85.5% given an ICER threshold of £50,000 per QALY.

Figure 44. Cost-effectiveness acceptability curves for blinatumomab and SoC maintenance therapy

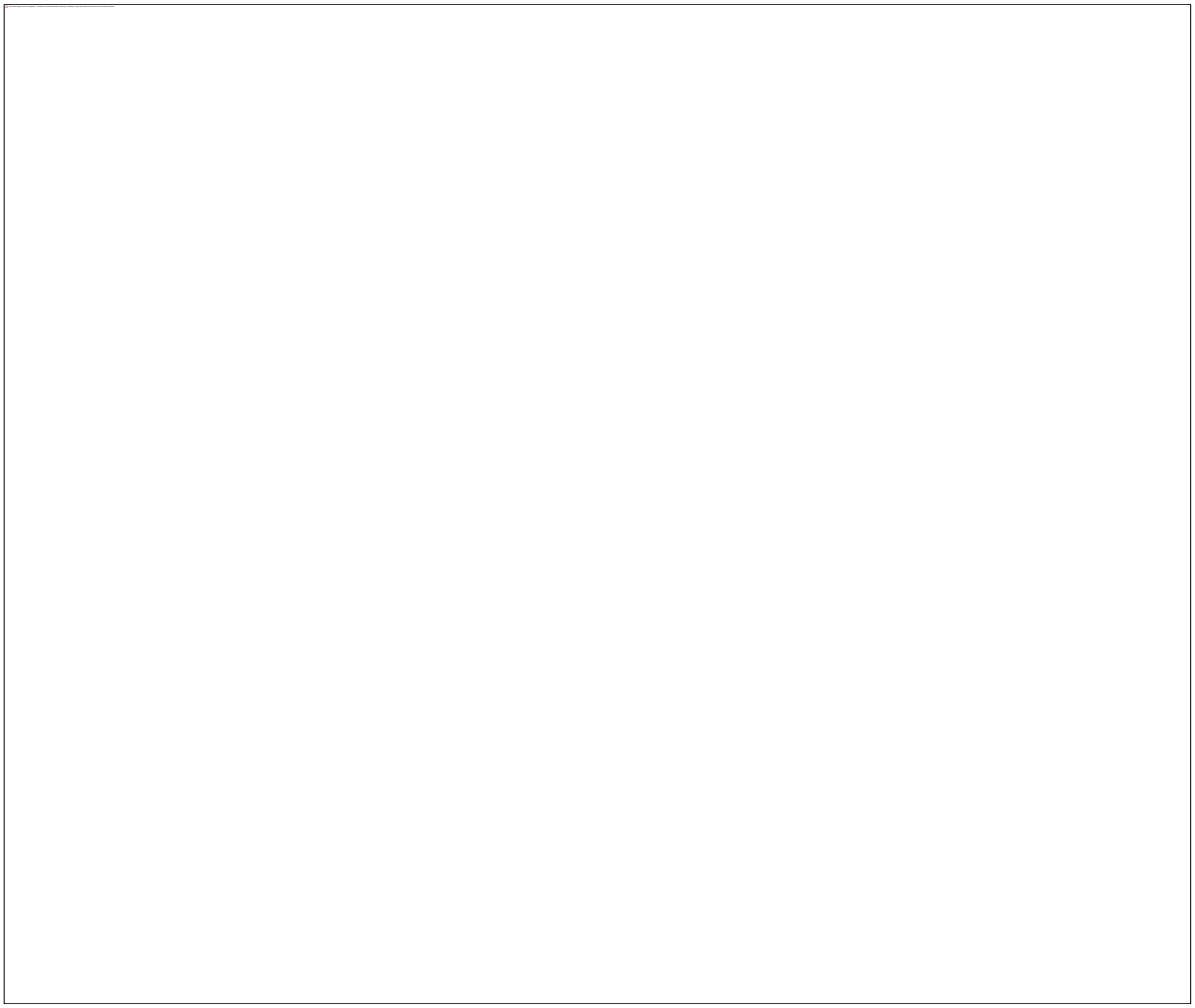


Abbreviations: SoC: standard of care; ICER: incremental cost-effectiveness ratio; WTP: willingness to pay threshold; QALY: quality-adjusted life year.

B.3.8.2 Deterministic sensitivity analysis

A tornado chart for the ICER for blinatumomab vs. SoC is shown in Figure 45. Changes in the proportion of blinatumomab patients receiving HSCT had a relatively large effect on the ICER, which varies from £16,408 to £44,322 per QALY gained as this parameter is varied across its 95% CI (per 6 months, base-case = 14.1%, 95% CI: 7.43% to 22.57%). The model was also relatively sensitive to the parameters relating to the duration of treatment with blinatumomab, as seen by varying the proportion starting and completing treatment, with the ICER varying from £23,260 to £34,101 per QALY gained as these parameters were varied simultaneously across their 95% CIs.

Figure 45. Tornado diagram of ICER of blinatumomab versus SoC



Abbreviations: ICER: incremental cost-effectiveness ratio; SoC: standard of care; OS: overall survival; Allo-SCT: allogeneic stem cell transplantation; IP: inpatient; MRD: minimal residual disease' RFS: relapse-free survival; OP: outpatient; Tx: treatment.

B.3.8.3 Scenario analysis

A description of the various scenario analyses is provided in Table 61.

Table 61. Description of scenario analyses

No.	Description	Base-case setting	Scenario setting	Justification
1	ATE weights	ATT weights	Utilities, MRD response rates, age, proportion male, duration of therapy, RFS distribution, OS distribution, probability RFS event is death, all with ATE rather than ATT weights.	Alternative methodology as per NICE DSU TSD 17 using ATE weights explored
2	Alternative Extrapolation Methods	RFS Gompertz (U), OS Lognormal Mix (Cure)	RFS and OS distributions changed to restricted Gompertz and unrestricted Weibull non-mixture cure, respectively.	Restricted Gompertz was the best-fitting RFS distribution based on the fit criteria used for distribution selection. The unrestricted Weibull non-mixture cure distribution was the best-fitting OS distribution that was compatible with the restricted Gompertz, i.e. RFS never exceeded OS. This combination presents a more favourable scenario.
3		RFS Gompertz (U), OS Lognormal Mix (Cure)	RFS and OS distributions changed to restricted RCS log-logistic and restricted RCS Weibull, respectively.	The RCS log-logistic was the third-best fitting distribution for RFS (the second was used for the base-case) based on the fit criteria used for distribution selection. The restricted RCS Weibull distribution was the best fitting OS distribution that was compatible with the selected RFS distribution, i.e. RFS never exceeded OS, and the second-best OS distribution overall. This combination presents a less favourable scenario.
4	2-fold increase in long-term excess mortality	4-fold increase in long-term excess mortality	Long-term excess mortality set to 2 (scenario 4) and 6 (scenario 5).	The base-case assumed a minimum of a 4-fold increase in mortality versus general population based on an analysis of the long-term consequences of allogeneic HSCT conducted by Martin et al.[78] We evaluated the sensitivity of the model to this assumption by increasing and decreasing this estimate by 50%.
5	6-fold increase in long-term excess mortality			

No.	Description	Base-case setting	Scenario setting	Justification
6	Duration of benefits = 60 months	In the base-case, RFS and OS were modelled based on parametric survival distributions fit to survival data from BLAST and the historical control, combined with age- and sex-matched general population mortality adjusted for excess risk of death due to exposure to radiotherapy, chemotherapy, and HSCT. This approach was assumed to accurately represent the long-term benefits of blinatumomab on survival. Based on this approach, the HR for OS for blinatumomab versus SoC reached a nadir of approximately 0.37 at 8 years and was equal to approximately 1.0 by 11 years. Hence this approach implicitly limits the duration of benefit on OS to 11 years.	Duration of benefits set to 60 months.	While the base-case assumption implicitly limits the duration of benefits of blinatumomab on survival, this scenario was generated to investigate the impact of explicitly limiting the duration of benefit to 60 months. 60 months was chosen as the point when patients are considered "cured" and therefore no longer under the influence of blinatumomab.
7	Inpatient costs with on-treatment inpatient days from BLAST	4 inpatient days cycle 1, 2 inpatient days cycle 2, 0 inpatient days thereafter, based on the NICE guidance TA450 for R/R Ph- B-cell precursor ALL[75]	8.8 inpatient days' cycle 1, 5.4 inpatient days' cycle 2, 4.2 inpatient days' cycle 3, 3.8 inpatient days' cycle 4.	The base-case uses the number of inpatient days outlined in the NICE guidance TA450 for blinatumomab for R/R ALL.[75] For sensitivity, we generated results first using the number of inpatient days observed in the BLAST trial for the CR1 population and then based on the number of inpatient days in the proposed EMA SmPC for blinatumomab MRD indication.
8	Inpatient costs with on-treatment inpatient days from blinatumomab label		3 inpatient days' cycle 1, 2 inpatient days in each subsequent cycle (cycles 2-4).	

No.	Description	Base-case setting	Scenario setting	Justification
9	Blinatumomab RFS events that are deaths	47.1%	20%	Because the relatively high proportion of RFS events that were deaths for blinatumomab may reflect incomplete capture of relapses after transplant in BLAST, a scenario analysis was conducted assuming the proportion of RFS that were deaths was only 20%.
10	HRU data from online survey	In the base-case, HRU data were based on results of face-to-face interviews of two UK clinicians: • Inpatient days MRD+: 1.75 • Inpatient days MRD-: 0.06 • Visits to haematologist, MRD+: 2.00 • Visits to haematologist, MRD-: 1.50 • Visits to radiologist, MRD+: 0.42 • Visits to radiologist, MRD-: 0.25 • Visits to physician, MRD+: 0.75 • Visits to physician, MRD-: 0.42 • Other visits, MRD+: 0.50 • Other visits, MRD-: 0.25	In the scenario analysis, HRU was based on results of the online survey of 20 UK clinicians: • Inpatient days MRD+: 3.10 • Inpatient days MRD-: 2.33 • Visits to haematologist, MRD+: 1.17 • Visits to haematologist, MRD-: 0.92 • Visits to radiologist, MRD+: 0.25 • Visits to radiologist, MRD-: 0.08 • Visits to physician, MRD+: 0.83 • Visits to physician, MRD-: 0.50 • Other visits, MRD+: 0.25 • Other visits, MRD-: 0.25	To investigate the impact of alternative data source for HRU associated with MRD In the base-case, inpatient and outpatient healthcare resource utilisation (HRU) by MRD response was based on results of face-to-face interviews of two UK experts – this approach was considered appropriate given the rare and complex nature of this disease area. Nevertheless, a follow-up, larger multinational online survey that was also conducted to gather more information on patterns of testing for MRD response. The results for the online survey were considered only in a scenario analysis as despite the increased sample size, the distribution of results received suggested that many physicians participating in the online survey did not adequately understand the questions, thus this likely reflected a less accurate estimate of the resource impact.
11	Cumulative probability of pre-relapse HSCT identical for	The cumulative probability of pre-relapse HSCT for CR1 population of BLAST trial was 72.6%. The six-month	The cumulative probability of pre-relapse HSCT for patients in the historical control study was 38.4%. The six-month probability for	A high rate of HSCT was observed in the BLAST trial, which might not be accurately reflecting the UK clinical practice, given that a large proportion of

No.	Description	Base-case setting	Scenario setting	Justification
	blinatumomab as for SoC	probability for months 1–48 was estimated to be 14.15%	months 1–48 that yielded this value at 48 months for blinatumomab patients was 7.47%	the patients in BLAST are from Germany. This scenario was run to investigate results using an HSCT rate equal to that observed in the historical control study.
12	ALL-related costs applied indefinitely	ALL-related costs applied to 60 months	Time when ALL-related costs not applied set to infinity.	To investigate the sensitivity of the model to assumptions regarding ALL-related costs.
13	0% MRD response rate for SoC	8% MRD response rate for SoC	SoC MRD response rate set to 0%.	To investigate other reasonable assumptions about the MRD response rate for SoC.
14	15% MRD response rate for SoC		SoC MRD response rate set to 15%.	
15	No disutility for long-term survivors	0.02 disutility for long-term survivors	Set disutility for long-term survivors to 0.	To investigate other reasonable assumptions regarding disutility for long-term survivors.
16	0.04 disutility for long-term survivors		Set disutility for long-term survivors to 0.04.	
17	SoC RFS utility equal to blinatumomab off-treatment RFS utility	Utility during RFS for patients receiving SoC was estimated to be 0.806 based on the estimated utility value from the GLM/GEE regression analysis of EQ-5D utility values in BLAST for patients who were off treatment, in haematological relapse, and assuming 8% MRD response	Utility during RFS for patients receiving SoC was set to 0.842 based on the estimated utility value from the GLM/GEE regression analysis of EQ-5D utility values in BLAST for patients who were off treatment, in haematological relapse, and assuming the same MRD response as blinatumomab (83.5%)	To address any the impact of base-case assumption that blinatumomab patients having a higher utility during RFS than SoC patients as a consequence of higher rate of MRD response.
18	Use ALL-related utilities and costs only to 36 months	ALL-related utilities and costs used up to 60 months	Set times when pre-relapse other inpatient/outpatient, post-relapse other inpatient/outpatient, salvage, and terminal care costs no longer applied, as well as the time beyond which general population utilities are used and the terminal	To investigate the sensitivity of the model to the time when ALL-related costs and utilities are no longer applied, i.e., patients are cured after 36 months or 48 months.
19	Use ALL-related utilities and costs only to 48 months			

No.	Description	Base-case setting	Scenario setting	Justification
			decrement is no longer applied to 36 and 48 months, respectively.	
20	Model timeframe = 30 years	Model timeframe = 50 years	Model timeframe set to 30 and 60 years, respectively	To investigate the impact on model results of varying the model timeframe.
21	Model timeframe = 60 years			
22	Annual discount rate for costs and QALYs = 1.5%	Discount rates for costs and effectiveness are 3%	Discount rates for costs and effectiveness set to 1.5%.	To investigate the alternative discount rate suggested by the NICE Guide to Technology Appraisal.[86]
23	Limitations relating to generalisability of SoC arm to current practice	ATT-weighted analysis of historical cohort analysis	RFS and OS survival distribution based on the ATT analysis of the historical cohort study is adjusted upwards by a factor of 15%.	To account for potential limitations in the generalisability of historical data to current clinical practice. Although clinical expert opinion concluded that standard of care has not meaningfully progressed in the last decade, such an analysis may represent a 'worst-case' estimate of the relative efficacy.
24	Blinatumomab OS cure fraction = midpoint OS cure fractions	Blinatumomab cure fraction estimated from OS Lognormal Mixture Cure model	Set cure fraction = to midpoint between estimated SoC and blinatumomab cure fractions (effective 50% reduction in benefit)	To investigate impact on model to variations in the projected cure fraction

Abbreviations: HSCT: haematopoietic stem cell transplant; HSCT: haematologic stem cell transplant; MRD: minimal residual disease; QALY: quality-adjusted life-year; RFS: relapse-free survival; SoC: standard of care; OS: overall survival.

Results of scenario analyses are presented in Table 62.

Table 62. Results of scenario analyses

#	Scenario	Blinatumomab			SoC			Blinatumomab vs. SoC			
		Cost (£)	Life-Years	QALYs	Cost (£)	Life-Years	QALYs	Cost (£)	Life-Years	QALYs	ICER (£)
	Base case	[redacted]	9.09	7.10	[redacted]	5.51	4.14	84,259	3.58	2.95	28,524
1	ATE weights	[redacted]	8.70	7.01	[redacted]	5.99	4.62	81,370	2.71	2.39	33,999
2	Alternative extrapolation methods Unfavourable - RFS RCS Log-Logistic (R), OS RCS Weibull (R)	[redacted]	9.47	7.40	[redacted]	5.44	4.09	83,064	4.02	3.31	25,081
3		[redacted]	8.57	6.70	[redacted]	5.26	3.96	83,874	3.30	2.74	30,647
4	2-fold increase long-term excess mortality	[redacted]	10.03	7.80	[redacted]	5.93	4.46	84,300	4.10	3.35	25,199
5	6-fold increase long-term excess mortality	[redacted]	8.46	6.63	[redacted]	5.23	3.94	84,234	3.23	2.69	31,274
6	Duration of benefits = 60 months	[redacted]	8.42	6.58	[redacted]	5.51	4.14	84,263	2.91	2.44	34,559
7	IP costs with on-Tx IP days from BLAST	[redacted]	9.09	7.10	[redacted]	5.51	4.14	89,235	3.58	2.95	30,209
8	IP costs with on-Tx IP days from Blincyto® label	[redacted]	9.09	7.10	[redacted]	5.51	4.14	84,405	3.58	2.95	28,574
9	23.55% of blinatumomab RFS events are deaths	[redacted]	9.09	7.09	[redacted]	5.51	4.14	90,548	3.58	2.95	30,698

10	HRU data from online survey	[redacted]	9.09	7.10	[redacted]	5.51	4.14	105,376	3.58	2.95	35,673
11	Cumulative probability of pre-relapse HSCT same for blinatumomab as for SoC	[redacted]	9.09	7.14	[redacted]	5.51	4.14	49,403	3.58	3.00	16,479
12	ALL-related costs applied to end of model time horizon	[redacted]	9.09	7.10	[redacted]	5.51	4.14	80,302	3.58	2.95	27,185
13	0% MRD response rate for SoC	[redacted]	9.09	7.10	[redacted]	5.51	4.14	82,537	3.58	2.96	27,892
14	15% MRD response rate for SoC	[redacted]	9.09	7.10	[redacted]	5.51	4.15	85,766	3.58	2.95	29,080
15	No disutility for long-term survivors	[redacted]	9.09	7.22	[redacted]	5.51	4.21	84,259	3.58	3.01	27,979
16	0.04 disutility for long-term survivors	[redacted]	9.09	6.98	[redacted]	5.51	4.08	84,259	3.58	2.90	29,091
17	SoC RFS utility = blinatumomab off-Tx RFS utility	[redacted]	9.09	7.10	[redacted]	5.51	4.16	84,259	3.58	2.93	28,722
18	ALL-related utilities and costs only to 36 months	[redacted]	9.09	7.11	[redacted]	5.51	4.19	87,100	3.58	2.92	29,866
19	ALL-related utilities and costs only to 48 months	[redacted]	9.09	7.10	[redacted]	5.51	4.17	85,364	3.58	2.94	29,056

20	Model timeframe = 30 y	[redacted]	8.78	6.88	[redacted]	5.35	4.03	84,126	3.42	2.85	29,552
21	Model timeframe = 60 y	[redacted]	9.09	7.10	[redacted]	5.51	4.14	84,259	3.58	2.95	28,524
22	Annual discount rate for costs and QALYs=1.5%	[redacted]	11.27	8.77	[redacted]	6.65	5.01	85,119	4.62	3.76	22,639
23	Limitations relating to generalisability of SoC arm to current practice	[redacted]	9.09	7.10	[redacted]	6.48	4.88	80,202	2.61	2.22	36,163
24	Blinatumomab OS cure fraction = midpoint OS cure fractions	[redacted]	7.34	5.75	[redacted]	5.51	4.14	78,918	1.83	1.61	49,101

Abbreviations: SoC: standard of care; LY: life years; QALY: quality-adjusted life year; ATE: average treatment effect; RFS: relapse-free survival; R: restricted; OS: overall survival; U: unrestricted; Tx: treatment; HRU: healthcare costs and resource use; HSCT: haematopoietic stem cell transplantation; ALL: acute lymphoblastic leukaemia

The impact of the key scenario analyses are discussed in more detail below.

The first scenario examined the cost-effectiveness of blinatumomab versus SoC using ATE rather than ATT weighting. ATE weights were applied to the RFS and OS survival distributions, utilities, duration of therapy, mean starting age, mean proportion of male patients, and mean body surface area (BSA) (see Appendix L and Appendix P). Cost effectiveness of blinatumomab is somewhat less favourable using the ATE weights, yielding an ICER of £33,999.

As outlined in the curve fitting section (see Section B.3.3), the models selected for the base-case were selected based on fit statistics, visual fit, and consistency of RFS and OS projections. Other survival distributions that were not selected but still performed well are presented in scenarios 2 and 3 (see Section B.3.3). Of the parametric cure models, we decided to use the more conservative of the best-fitting options as the base-case. Scenario 2 presents a more favourable selection whereas scenario 4 presents a less favourable approach. ICERs for these scenarios were £25,081 and £30,647 for scenarios 3 and 4, respectively.

The base-case uses estimates of HRU for follow-up and monitoring based on face-to-face interviews of 2 UK clinicians. In Scenario 10 HRU data from the online survey of 20 UK clinicians was used instead. The projected mean number of inpatient days was substantially greater, and the difference in mean inpatient days for MRD+ versus MRD- patients was substantially less, based on the online survey data versus the face-to-face interviews. Use of the online survey data therefore increased the ICER to £35,673 per QALY gained. However, as discussed in Section B.3.5.4, the HRU costs based on in-deth interviews were considered to more accurately reflect the true resource implications despite the smaller sample size.

In the base-case, the probability of allogeneic HSCT pre-relapse was estimated to be greater in patients receiving blinatumomab compared with SoC. In Scenario 12, the probability of allogeneic HSCT with blinatumomab was calibrated so that the cumulative probability of pre-relapse HSCT is the same for blinatumomab as for SoC. Because LYs and QALYs are estimated independently of the rate of HSCT, changes in this parameter only impact the expected costs. Given the high cost of HSCT, setting the cumulative probabilities of HSCT to be the same for blinatumomab and SoC reduced the ICER considerably, to £16,479 per QALY gained.

Finally, a further scenario analysis were conducted to explore limitations associated with estimating comparative effectiveness due to the single-arm BLAST. In scenario 23, the OS survival estimated for SoC was revised upwards by 15% (HR 0.85) to assess potential impact on the underestimation of survival based on historical data – the HR was selected based on an analysis presented in the NICE appraisal for blinatumamb for the treatment of R/R ALL where a historical control study was shown to underestimate survival when compared to the pivotal phase 3 study (TOWER). The resulting ICER increased to £36,163 per QALY gained. Nevertheless, clinical expert consistently concluded that standard of care has not meaningfully progressed in the last decade, therefore such an analysis may represent a ‘worst-case’ estimate of the relative efficacy. Furthermore, in the scenario analysis assessing the impact of varying the estimated cure fraction, the ICER remained cost-effective (£49,101) despite assuming a 50% relative reduction in benefit from the base case analysis

B.3.8.4 Summary of sensitivity analyses results

Results of the PSA indicate that there is 85.5% probability that blinatumomab is preferred to SoC at a WTP threshold of £50,000 per QALY. From the deterministic sensitivity analyses, changes in the proportion of blinatumomab patients receiving HSCT had the most impact on the ICER, which varies from £16,408 to £44,322 per QALY gained as this parameter is varied across its 95% CI (per 6 months, base-case = 14.1%, 95% CI: 7.43% to 22.57%). In all scenario analyses, the ICER for blinatumomab was consistently below the threshold of £50,000 per QALY for end of life medicines. The ICER was least favourable in the scenarios using the HRU data from the online survey (£35,673 per QALY), when the duration of benefits of blinatumomab on RFS and OS were limited to 60 months (£34,559 per QALY), and in the highly conservative analysis where the OS of the SoC arm is improved by 15% (£36,163 per QALY). Furthermore, in the scenario analysis assessing the impact of varying the estimated cure fraction, the ICER remained cost-effective (£49,101) despite assuming a 50% relative reduction in benefit from the base case analysis.

Importantly, in the key scenario analysis where blinatumomab was considered to be used as a salvage treatment in the Soc arm (to better reflect clinical practice), the ICER was significantly more favourable than in the base-case (£17,522 per QALY).

B.3.9 *Subgroup analysis*

No subgroup analyses were conducted.

B.3.10 *Validation*

B.3.10.1 Validation of cost-effectiveness analysis

We conducted a validation of the Microsoft Excel workbook model used in the economic evaluation. The validation consisted of a series of procedures to check the general modelling approach, model calculations, model functionality, and model inputs. The procedures employed to validate the model are summarised in Table 63.

Table 63. Summary of validation procedures

No.	Step
1	Exploratory tests of Model calculations
2	Identify unused named ranges;
3	Check that there are no links to other workbooks or external files
4	Test Model control objects (buttons, etc.) for functionality
5	Check that “load/save” works correctly
6	Identify overly complex/difficult to parse formulas
7	Identify #REF, #NUM, and #NA errors
8	Identify hard-coded values within formulas
9	Identify inconsistencies in formulas across contiguous ranges

10	Identify unused calculations
11	Check that all input values are appropriately referenced
12	Check index/lookup functions for offset errors
13	Check that x- and y-axis ranges on Model charts change as results change
14	Generate results using extreme values/edge cases for Model inputs and check Model results for errors or anomalous findings
15	Generate sensitivity analyses using the Model and check the results of these analyses against priors (e.g., increasing medication costs increases costs and ICER but not LYs and QALYs; increasing utilities increases QALYs but not LYs or costs, increasing mortality decreases LYs, QALYs and costs)
16	Generate probabilistic sensitivity analyses and check the results against those based on base-case point estimates (e.g., check that mean incremental costs from PSA equals incremental cost for base-case; check correlation of study outcomes)
17	Check the Model inputs against source documents
18	Check Model formatting (e.g., inputs one colour fill, results a different colour fill)
19	Check that Model is free of spelling and grammar errors
20	Check that discounting is applied appropriately

Abbreviations: ICER: incremental cost-effectiveness ratio; LY: life year; QALY: quality-adjusted life year.

B.3.10.2 Exploratory test

Exploratory testing is a form of *ad hoc* testing, whereby an experienced modeller probes the correctness of the model using different heuristics and edge cases or extreme values.[110-112] The following exploratory tests were performed:

- Discounting of LYs and QALYs was checked by setting discount rates for effectiveness to 0% and confirming that undiscounted and discounted LYs and QALYs were equal.
- Utility values were checked by setting all utility values to 1.0 and all disutilities to zero and confirming that LYs were equal to QALYs.
- Survival calculations were checked by setting RFS and OS to a custom distribution with RFS and OS equal to 100% for the entire model time horizon, setting model start age to 50 years and model sex to 100% male, and excess mortality to zero, and confirming that LYs were equal to UK general population life expectancy for males at 50 years of age (31.1 years in model versus 31.18 years in UK 2013-2015 lifetable).
- Survival calculations were checked by setting the RFS and OS to a custom distribution with RFS and OS equal to general population mortality for a 50-year-old male, setting model start age to 50 years, sex to 100% male, excess mortality to zero, and confirming that LYs are consistent with UK general population expectancy for males at 50 years of age with and without the use of the general population mortality as a floor.
- The calculation of the costs of blinatumomab medication was checked by generating a one-way sensitivity analysis on the unit cost of blinatumomab and confirming that changes in unit cost of blinatumomab results in linear changes in the ICER.

- Average utility values by state were back calculated from the results by dividing QALYs per state by the corresponding LYs and comparing results with model inputs for utility values to ensure consistency.
- Average blinatumomab medication costs per cycle were back calculated from the results by dividing the expected blinatumomab medication cost from the results by the average number of vials per patent (based on the percentage starting and completing each cycle and the number of vials per cycle), and comparing results with model inputs for medication costs to ensure consistency.
- Average costs of allogeneic HSCT per patient receiving HSCT pre-relapse were back calculated from the results by setting the discount rate for costs to zero, then dividing the expected costs of pre-relapse HSCT in the results by the model inputs for the cumulative probabilities of receiving HSCT pre-relapse, and comparing results with model inputs for costs of HSCT to ensure consistency.
- Average costs of salvage therapy per patient receiving salvage therapy was back calculated from the results by setting the discount rate for costs to zero, then dividing the expected costs of salvage therapy in the results by the cumulative proportion of patients relapsing, and comparing results with model inputs for costs of salvage therapy to ensure consistency.
- Average annual pre- and post-relapse inpatient and outpatient costs were back calculated from the results by setting the model timeframe to 5 years (to correspond to the period over which these costs are accrued in the base-case) and dividing the (discounted) expected pre- and post-relapse inpatient and outpatient costs in the results by the (discounted) expected pre- and post-relapse life expectancy (in years). Back-calculated annual costs were compared with annual costs derived directly from the model inputs to ensure consistency.

None of these tests identified any potential errors in the model calculations.

B.3.10.3 Named ranges

The named ranges in the model were checked to identify any ranges that were unused, referenced external workbooks, or had missing references. No named ranges with errors were identified. No named ranges were linked to any external workbooks.

B.3.10.4 Control objects

The control objects in the model including menu items, radio buttons, buttons, and drop-down options were checked to make sure that they work correctly:

- The “load/save inputs” feature was found to work appropriately: new scenarios were created, modified, and deleted as needed.
- All source boxes were tested and worked correctly.
- All dropdown menus were tested and worked correctly.
- The option to hide calculations on the PSA input sheet worked properly.

B.3.10.5 Formulas

The formulas used in the model were checked to identify the following:

1. Long and complicated formulas that are hard to evaluate
2. Formulas that returned errors (e.g. #N/A, #VALUE, #NUM, #REF!)
3. Hardcoded values in formulas
4. Inconsistencies in formulas across contiguous ranges
5. Unused calculations
6. Unused input values
7. Offset errors in index/lookup functions

No formulas were found to be unnecessarily long or complicated. The model was searched for "#" to identify any #N/A, #VALUE, #NUM, or #REF! errors; no such errors were identified. The only hardcoded values included model formulas were required for model calculations (e.g., LYs per model cycle entered as "7/365"). No inconsistent formulas across contiguous ranges were identified. No unused calculations were identified. No offset errors were identified.

B.3.10.6 Charts

The charts in the model were checked to make sure that the data used by graphs was appropriate, and that all the graphs were populated. No issues were identified with the x- and y-axes of charts.

B.3.10.7 Extreme values

To ensure the model generated expected results when extreme values were used as inputs, several analyses were conducted using extreme values for selected model inputs. All results met expectations.

B.3.10.8 Sensitivity analyses

To test whether the model generates results consistent with expectations, we ran the deterministic sensitivity analyses for the base-case (see Section B.3.8.2). We then calculated the difference in life years, QALYs, and costs for each sensitivity analysis versus the base-case (Table 64). For each treatment and scenario, we then specified our expectations for how the cost, life years, and QALYs, would be expected to change (increase, decrease, or no change) for the given change in the parameter value. We then compared the expected change with the actual change. Most results were consistent with expectations with two exceptions. Although it would seem that a healthier population would incur fewer costs, actually costs decrease with the lower bound of the OS cure fraction, for both blinatumomab and SoC. This is because there is a higher probability of receiving HSCT pre-relapse rather than post-relapse and since there is a lower proportion of patients in pre-relapse in these sensitivity analyses, fewer patients receive HSCT, and so incur fewer costs.

Table 64. Deterministic sensitivity analysis: difference vs. base-case (low value), expected vs. actual results

Difference versus Base-case	Expected Results (Up, Down, No Change)						Matches Expected Results					
	Blinatumomab			SoC			Blinatumomab			SoC		
	Costs	LYs	QALYs	Costs	LYs	QALYs	Costs	LYs	QALYs	Costs	LYs	QALYs
Blin MRD Response (95% CI)	+	0	-	+	0	-	✓	✓	✓	✓	✓	✓
Blin RFS Hazard Ratio (95% CI)	+	0	-	0	0	0	✓	✓	✓	✓	✓	✓
Blin RFS Shape Difference (95% CI)	+	0	-	0	0	0	✓	✓	✓	✓	✓	✓
Blin OS Cure Fraction Difference (95% CI)	+	-	-	0	0	0	✘	✓	✓	✓	✓	✓
Blin IP On-Tx Cost (± 50%)	-	0	0	0	0	0	✓	✓	✓	✓	✓	✓
Blin OP On-Tx Cost (± 50%)	-	0	0	0	0	0	✓	✓	✓	✓	✓	✓
Blin Duration of Therapy (95% CI)	-	0	+	0	0	0	✓	✓	✓	✓	✓	✓
Proportion Blin Receiving HSCT (95% CI)	-	0	0	0	0	0	✓	✓	✓	✓	✓	✓
Proportion SoC Receiving HSCT (95% CI)	0	0	0	-	0	0	✓	✓	✓	✓	✓	✓
Proportion Post-Relapse HSCT (95% CI)	-	0	0	-	0	0	✓	✓	✓	✓	✓	✓
HSCT Costs (± 50%)	-	0	0	-	0	0	✓	✓	✓	✓	✓	✓

Multi-Agent Chemo Salvage Costs (± 50%)	-	0	0	-	0	0	✓	✓	✓	✓	✓	✓
Other IP Costs (± 50%)	-	0	0	-	0	0	✓	✓	✓	✓	✓	✓
Other OP Visits Costs (± 50%)	-	0	0	-	0	0	✓	✓	✓	✓	✓	✓
SoC OP Visits Costs (± 50%)	0	0	0	-	0	0	✓	✓	✓	✓	✓	✓
Post-Relapse Utility (95% CI)	0	0	-	0	0	-	✓	✓	✓	✓	✓	✓
Baseline Mean Utility (95% CI)	0	0	-	0	0	-	✓	✓	✓	✓	✓	✓
Intercept Utility (95% CI)	0	-	-	0	-	-	✓	✓	✓	✓	✓	✓
MRD Response Utility Coefficient (95% CI)	0	-	-	0	-	-	✓	✓	✓	✓	✓	✓
Blin Prob. RFS Event Death (95% CI)	+	0	0	0	0	0	✓	✓	✓	✓	✓	✓
SoC Prob. RFS Event Death (95% CI)	0	+	0	0	0	0	✓	✓	✓	✓	✓	✓
Off-Tx Relapse-Free Utility Coefficient (95% CI)	0	-	-	0	-	-	✓	✓	✓	✓	✓	✓
Terminal Utility Decrement (95% CI)	0	-	-	0	-	-	✓	✓	✓	✓	✓	✓
SoC OS Cure Fraction (95% CI)	0	0	0	+	-	-	✘	✓	✓	✓	✓	✓
Utility Decrements for HSCT (95% CI)	0	0	0	0	-	-	✓	✓	✓	✓	✓	✓

Abbreviations: SoC: standard of care; LY: life year; QALY: quality-adjusted life year; MRD: minimal residual disease; RFS: relapse-free survival; OS: overall survival; CI: confidence interval; IP: inpatient; OP: outpatient; Tx: treatment; HSCT: haematopoietic stem cell transplantation.

B.3.10.9 Scenario analyses

Twenty-two scenarios were generated to test assumptions in the model. Table 65 compares the expected result of the scenarios (whether values should increase, decrease, or remain the same) to the actual results. The scenarios using ATE weights and different survival distributions are excluded from this analysis because it is initially unclear how the results would be affected. Somewhat unexpectedly, decreasing excess mortality (slightly) decreased costs for SoC while increasing costs for blinatumomab. This is likely due to the relatively large proportion of SoC costs that are associated with terminal care, as the reduction in terminal costs offsets the increase in post-relapse costs when excess mortality is reduced. All other scenarios yielded results that were consistent with expectations.

Table 65. Scenario analysis: difference vs base-case, expected vs actual results

Scenario	Expected Results (Up, Down, No Change)						Matches Expected Results					
	Blinatumomab			SoC			Blinatumomab			SoC		
	Costs	LYs	QALYs	Costs	LYs	QALYs	Costs	LYs	QALYs	Costs	LYs	QALYs
2-fold long-term excess mortality	+	+	+	-	+	+	✓	✓	✓	x	✓	✓
6-fold long-term excess mortality	-	-	-	+	-	-	✓	✓	✓	x	✓	✓
Duration of benefits = 60 months	+	-	-	0	0	0	✓	✓	✓	✓	✓	✓
IP costs with on-Tx IP days from BLAST	+	0	0	0	0	0	✓	✓	✓	✓	✓	✓
IP costs with on-Tx IP days from Blincyto label	+	0	0	0	0	0	✓	✓	✓	✓	✓	✓
Use updated HRU data	+	0	0	+	0	0	✓	✓	✓	✓	✓	✓
Cumulative probability of pre-relapse HSCT same for blinatumomab as for SoC	-	0	0	0	0	0	✓	✓	✓	✓	✓	✓
Blinatumomab salvage Tx for SoC	0	0	0	+	0	+	✓	✓	✓	✓	✓	✓
ALL-related costs applied indefinitely	+	0	0	+	0	0	✓	✓	✓	✓	✓	✓
0% MRD response rate for SoC	0	0	0	+	0	-	✓	✓	✓	✓	✓	✓
15% MRD response rate for SoC	0	0	0	-	0	+	✓	✓	✓	✓	✓	✓
No disutility for long-term survivors	0	0	+	0	0	+	✓	✓	✓	✓	✓	✓
0.04 disutility for long-term survivors	0	0	-	0	0	-	✓	✓	✓	✓	✓	✓

SoC RFS utility = blinatumomab off-Tx RFS utility	0	0	0	0	0	+	✓	✓	✓	✓	✓	✓
Use ALL-related utilities and costs only to 36 months	-	0	+	-	0	+	✓	✓	✓	✓	✓	✓
Use ALL-related utilities and costs only to 48 months	-	0	+	-	0	+	✓	✓	✓	✓	✓	✓
Model timeframe = 30 years	-	-	-	-	-	-	✓	✓	✓	✓	✓	✓
Model timeframe = 60 years	0	0	0	0	0	0	✓	✓	✓	✓	✓	✓
Annual discount rate for costs and QALYs = 1.5%	+	+	+	+	+	+	✓	✓	✓	✓	✓	✓

Abbreviations: SoC: standard of care; LY: life year; QALY: quality-adjusted life year; MRD: minimal residual disease; RFS: relapse-free survival; OS: overall survival; CI: confidence interval; IP: inpatient; OP: outpatient; Tx: treatment; HSCT: haematopoietic stem cell transplantation; HRU: healthcare resource use.

B.3.10.10 PSA results

The results generated by the PSA were compared with the results generated by the base-case to ensure that the mean costs, QALYs, and ICER generated by the PSA were similar to those generated by the model. The percent differences in the LYs, QALYs, and costs generated by the PSA versus the deterministic base-case results are shown in Table 66. The differences between the PSA and the deterministic base-case results were relatively modest (<5%). The ICER from the PSA (£29,673) was £1,149 (4.0%) greater than the ICER in the deterministic base-case (£28,524). These modest differences are not unexpected if costs and QALYs are nonlinear functions.

Table 66. Percent difference in results from PSA vs. deterministic base-case

	Life Years (Not Discounted)	QALYs (Discounted)	Cost (Discounted) (£)
Blinatumomab	-2.33%	-2.01%	0.14%
SoC	-1.29%	-0.84%	0.08%
Incremental	-3.77%	-3.65%	0.23%

Abbreviations: PSA: probabilistic sensitivity analyses; QALY: quality-adjusted life year; SoC: standard of care.

B.3.10.11 Model inputs

The inputs that are in the base-case of the model were checked against those reported in the corresponding source documents to ensure consistency. All inputs matched with the values reported in the source documents (see Table 67).

Table 67. Validation of inputs against external data sources

Parameter	Source	Correct?
Cost per inpatient stay (£)	2015-2016 NHS Reference Costs. Weighted average of number of stays for codes SA24G, SA24H, SA24J	✓
Cost per inpatient day (£)	2015-2016 NHS Reference Costs. Weighted average of number of stays over length of stay for codes SA24G, SA24H, SA24J	✓
Cost per visit to outpatient infusion centre (£)	2015-2016 NHS Reference Costs. Code SB15Z, "Deliver Subsequent Elements of a Chemotherapy Cycle"	✓
Cost per nurse visit (£)	2015-2016 NHS Reference Costs. Code N10AF, "Specialist Nursing, Cancer Related, Adult, Face to Face"	✓
Total cost of pump per patient (£)	Amgen	✓
Useful life expectancy of pump (years)	Amgen	✓
CSF prophylaxis dosage	BLAST CSR	✓
Cost of CSF prophylaxis drugs	eMit, 2015	✓
CSF prophylaxis administration cost	2015-2016 NHS Reference Costs. Code SB13Z, "Deliver More Complex Parenteral Chemotherapy at First Attendance"	✓

SoC maintenance chemotherapy dosage	UKALL14: A randomized trial for adults with newly diagnosed acute lymphoblastic leukaemia. NCT01085617	✓
Cost of SoC maintenance chemotherapy medication	eMit, 2015; Medicines Complete (BNF)	✓
SoC maintenance chemotherapy administration costs	2015-2016 NHS Reference Costs. Code SB14Z, "Deliver Complex Chemotherapy, included Prolonged Infusional Treatment, at First Attendance" and Code SB13Z, "Deliver More Complex Parenteral Chemotherapy at First Attendance"	✓
Probability of receiving HSCT post-relapse	Gokbuget N, Kelsh M, Chia V, et al. Blinatumomab vs historical standard therapy of adult relapsed/refractory acute lymphoblastic leukemia. Blood Cancer J. 2016; 6: e473.	✓
Cost of HSCT	Unrelated Donor Stem Cell Transplantation in the UK. Updated from 2012/13 costs to 2015/16 using PSSRU	✓
Cost of cyclosporine	Medicines Complete (BNF)	✓
Cost of salvage chemotherapy	TOWER UK NICE Submission	✓
Other inpatient stays/days by MRD response	Pilot HRU data	✓
Other outpatient visits by MRD response	Pilot HRU data	✓
Other outpatient visit costs	2015-2016 NHS Reference Costs. Code WF01A-303, "Non-Admitted Face to Face Attendance, Follow-Up, Clinical Haematology"; Code WF01A-812, "Non-Admitted Face to Face Attendance, Follow-Up, Diagnostic Imaging"; Code WF01A-370, "Non-Admitted Face to Face Attendance, Follow-Up, Medical Oncology". PSSRU, 2016 General practioner unit costs per patient contact lasting 9.22 minutes	✓
Terminal care cost	Addicott R, Dewar S. King`s Fund. Improving choice at end of life. A descriptive analysis of the impact and costs of the Marie Curie delivering choice programme in Lincolnshire.; Marie Curie Cancer Care. Understanding the cost of end of life care in different settings.; Adjusted to 2015/16 pounds using PSSRU	✓
Discounted incremental QALYs for receipt of blinatumomab vs. SoC as salvage therapy	TOWER UK NICE Submission	✓
General population utilities	Kind P, Hardman G, Macran S. UK Population Norms for EQ-5D. The University of York Centre for Health Economics.	✓
General population mortality	Office for National Statistics. England, Interim Life Tables, 1980-82 to 2013-2015. 2015.	✓

Abbreviations: NHS: National Health Service; CSR: clinical study report; CSF: cerebrospinal fluid; SoC: standard of care; HSCT: haematopoietic stem cell transplantation; BNF: British National Formulary; HRU: healthcare resource use; MRD: minimal residual disease; QALY: quality-adjusted life year.

B.3.10.12 Formatting

The formatting of the input and results cells was checked for consistency with what is described in the 'Introduction' sheet of the model. No errors in formatting were identified.

B.3.10.13 Spell check

A spell check was run on all the sheets to identify any spelling mistakes. No errors were identified.

B.3.10.14 Discounting

Discounting was applied to life years, QALYs and costs. Discounting was applied on an annual basis beginning one year after entry into the model. No errors in the calculation of the discount factor were identified.

B.3.11 *Interpretation and conclusions of economic evidence*

A Microsoft Excel-based partitioned-survival analysis model was used to evaluate the cost-effectiveness of blinatumomab versus SoC maintenance therapy in patients with MRD+ B-precursor ALL in haematological CR using propensity matched data from the BLAST and the historical control studies (MT103-203 and 20120148, respectively) and other sources. Based on this model, the ICER for blinatumomab versus SoC was estimated to be £28,524 per QALY gained. The mean ICER from the PSA (calculated as the ratio of the mean incremental costs to the mean incremental QALYs) was £29,673 per QALY gained. In the PSA, the probability that blinatumomab is preferred was estimated to be 85.5% given an ICER threshold of £50,000 per QALY. The ICER was substantially more favourable (£17,522) in the key scenario analysis where blinatumomab was considered as a salvage therapy for patients receiving SoC. Although the cost effectiveness evidence does not consider the Ph+ population or later remission states, due to the substantial unmet need across all sub-populations blinatumomab should be considered for use in alignment with its full anticipated marketing authorisation.

These results suggest that the use of blinatumomab in patients with MRD+ B-precursor ALL in haematological CR is a cost-effective use of healthcare resources given an ICER threshold of £50,000 per QALY. Furthermore, given that blinatumomab is indicated for a rare condition in a very small number of patients (85 per year) who have a huge unmet medical need and who stand to gain substantially from access to blinatumomab, this therapy meets many of the criteria for appraisal under the HST framework. Consequently, blinatumomab should be evaluated taking into account a wider range of criteria about the benefits and costs. [key ask]

In conclusion, blinatumomab is a cost-effective, highly innovative and well-tolerated therapy with demonstrable efficacy to achieve MRD negativity. As the first and only therapy specifically indicated for MRD+ BCP-ALL, blinatumomab represents a paradigm-shift and potentially curative treatment for patients with this rare and deadly cancer. **As such, we propose that blinatumomab is recommended for the treatment of MRD+ BCP-ALL in line with its anticipated licensed indication.**

References

1. Hoelzer D, Bassan R, Dombret H, et al. Acute lymphoblastic leukaemia in adult patients: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2016;27:v69-v82.
2. van Dongen JJ, van der Velden VH, Bruggemann M, et al. Minimal residual disease diagnostics in acute lymphoblastic leukemia: need for sensitive, fast, and standardized technologies. Blood 2015;125:3996-4009.
3. National Institute for Health and Care Excellence. Blinatumomab for acute lymphoblastic leukaemia for people with minimal residual disease activity in remission: final scope. 2017.
4. Nagorsen D, Bargou R, Ruttinger D, et al. Immunotherapy of lymphoma and leukemia with T-cell engaging BiTE antibody blinatumomab. Leuk Lymphoma 2009;50:886-91.
5. Gökbuget N, Dombret, H., Bonifacio, M., Reichle, A., Graux, C., Havelange, V., Buss, E. C., Faul, C., Bruggemann, M., Ganser, A., Stieglmaier, J., Wessels, H., Haddad, V., Zugmaier, G., Nagorsen, D., & Bargou, R. C. BLAST: A Confirmatory, Single-Arm, Phase 2 Study of Blinatumomab, a Bispecific T-Cell Engager (BiTE®) Antibody Construct, in Patients with Minimal Residual Disease B-Precursor Acute Lymphoblastic Leukemia (ALL). Blood 2014;124:379.
6. European Medicines Agency. Public summary of opinion on orphan designation: blinatumomab for the teratment of acute lymphoblastic leukaemia. 2014.
7. British National Formulary. Blinatumomab, 2017.
8. Amgen Inc. Summary of Product Characteristics: BLINCYTO 38.5 micrograms powder for concentrate and solution for infusion. 2017.
9. Pui CH, Robison LL, Look AT. Acute lymphoblastic leukaemia. Lancet 2008;371:1030-43.
10. Pui CH, Jeha S. New therapeutic strategies for the treatment of acute lymphoblastic leukaemia. Nat Rev Drug Discov 2007;6:149-65.
11. National Cancer Institute. Adult Acute Lymphoblastic Leukemia Treatment (PDQ®)–Health Professional Version, 2017.
12. Amgen Inc. ALL epidemiology estimates. 2017.
13. Fielding AK, Richards SM, Chopra R, et al. Outcome of 609 adults after relapse of acute lymphoblastic leukemia (ALL); an MRC UKALL12/ECOG 2993 study. Blood 2007;109:944-50.
14. Moorman AV, Chilton L, Wilkinson J, et al. A population-based cytogenetic study of adults with acute lymphoblastic leukemia. Blood 2010;115:206-14.
15. Khaled SK, Thomas SH, Forman SJ. Allogeneic hematopoietic cell transplantation for acute lymphoblastic leukemia in adults. Curr Opin Oncol 2012;24:182-90.
16. Thomas X, Boiron JM, Huguet F, et al. Outcome of treatment in adults with acute lymphoblastic leukemia: analysis of the LALA-94 trial. J Clin Oncol 2004;22:4075-86.
17. Moorman AV, Harrison CJ, Buck GA, et al. Karyotype is an independent prognostic factor in adult acute lymphoblastic leukemia (ALL): analysis of cytogenetic data from patients treated on the Medical Research Council (MRC) UKALLXII/Eastern Cooperative Oncology Group (ECOG) 2993 trial. Blood 2007;109:3189-97.
18. Oriol A, Vives S, Hernandez-Rivas JM, et al. Outcome after relapse of acute lymphoblastic leukemia in adult patients included in four consecutive risk-adapted trials by the PETHEMA Study Group. Haematologica 2010;95:589-596.
19. Gökbuget N, Kneba M, Raff T, et al. Adult patients with acute lymphoblastic leukemia and molecular failure display a poor prognosis and are candidates for stem cell transplantation and targeted therapies. Blood 2012;120:1868-1876.
20. Chiaretti S, Vitale A, Cazzaniga G, et al. Clinico-biological features of 5202 patients with acute lymphoblastic leukemia enrolled in the Italian AIEOP and GIMEMA protocols and stratified in age cohorts. Haematologica 2013;98:1702-10.
21. Campana D. Minimal residual disease in acute lymphoblastic leukemia. Hematology Am Soc Hematol Educ Program 2010;2010:7-12.
22. Hoelzer D, Thiel E, Loffler H, et al. Prognostic factors in a multicenter study for treatment of acute lymphoblastic leukemia in adults. Blood 1988;71:123-31.

23. Larson RA, Dodge RK, Burns CP, et al. A five-drug remission induction regimen with intensive consolidation for adults with acute lymphoblastic leukemia: cancer and leukemia group B study 8811. Blood 1995;85:2025-37.
24. Lussana F, Intermesoli T, Gianni F, et al. Achieving Molecular Remission before Allogeneic Stem Cell Transplantation in Adult Patients with Philadelphia Chromosome-Positive Acute Lymphoblastic Leukemia: Impact on Relapse and Long-Term Outcome. Biol Blood Marrow Transplant 2016;22:1983-1987.
25. Bassan R, Spinelli O, Oldani E, et al. Improved risk classification for risk-specific therapy based on the molecular study of minimal residual disease (MRD) in adult acute lymphoblastic leukemia (ALL). Blood 2009;113:4153-62.
26. Bruggemann M, Raff T, Flohr T, et al. Clinical significance of minimal residual disease quantification in adult patients with standard-risk acute lymphoblastic leukemia. Blood 2006;107:1116-23.
27. Thomas DA, Kantarjian H, Smith TL, et al. Primary refractory and relapsed adult acute lymphoblastic leukemia: characteristics, treatment results, and prognosis with salvage therapy. Cancer 1999;86:1216-30.
28. National Comprehensive Cancer Network. Practice Guidelines in Oncology, Acute Lymphoblastic Leukemia, 2016.
29. Bassan R, Hoelzer D. Modern therapy of acute lymphoblastic leukemia. J Clin Oncol 2011;29:532-43.
30. Cancer Research UK. Acute lymphoblastic leukaemia (ALL) incidence statistics, 2016.
31. Holowiecki J, Krawczyk-Kulis M, Giebel S, et al. Status of minimal residual disease after induction predicts outcome in both standard and high-risk Ph-negative adult acute lymphoblastic leukaemia. The Polish Adult Leukemia Group ALL 4-2002 MRD Study. Br J Haematol 2008;142:227-37.
32. Gokbuget N, Stanze D, Beck J, et al. Outcome of relapsed adult lymphoblastic leukemia depends on response to salvage chemotherapy, prognostic factors, and performance of stem cell transplantation. Blood 2012;120:2032-41.
33. Tavernier E, Boiron JM, Huguet F, et al. Outcome of treatment after first relapse in adults with acute lymphoblastic leukemia initially treated by the LALA-94 trial. Leukemia 2007;21:1907-14.
34. Berry DA, Zhou S, Higley H, et al. Association of Minimal Residual Disease With Clinical Outcome in Pediatric and Adult Acute Lymphoblastic Leukemia: A Meta-analysis. JAMA Oncol 2017:e170580.
35. Bar M, Wood BL, Radich JP, et al. Impact of minimal residual disease, detected by flow cytometry, on outcome of myeloablative hematopoietic cell transplantation for acute lymphoblastic leukemia. Leuk Res Treatment 2014;2014:421723.
36. Amgen Inc. TOWER PRO CSR: ALLSS baseline score. 2017.
37. Dombret H, Thomas X, Chevallier P, et al. Healthcare burden and reimbursement of hospitalization during chemotherapy for adults with Ph-negative relapsed or refractory B-cell precursor acute lymphoblastic leukemia in France: a retrospective chart review. J Med Econ 2016;19:1034-1039.
38. Pagano L, Oberti, M, Esposito, B, Reitan, J, Barlev, A, Barber, BL, Ferrara, F. Retrospective chart review of hospitalisations during chemotherapy for adult patients with Ph-negative B-precursor relapsed or refractory (R/R) acute lymphoblastic leukemia (ALL) in Italy. ISPOR 18th Annual European Congress 2015.
39. Graux C, Strens D, Hoefkens C, et al. Retrospective chart review of hospitalisations and costs associated with the treatment of adults with Philadelphia-negative (-) B-precursor relapsed or refractory (R/R) acute lymphoblastic leukemia (ALL) in Belgium. Belgian Hematological Society 31st Annual General Meeting 2016.
40. Kreuzer KA, Stuhlmann R, Lebioda A, et al. Hospitalisations among adult patients with Ph-negative B-precursor relapsed or refractory (R/R) acute lymphoblastic leukemia (ALL) receiving chemotherapy in Germany: a retrospective chart review. ISPOR 18th Annual European Congress 2015.

41. Boluda B, Rodrguez-Veiga R, Martinez-Cuadron D, et al. Time and cost of hospitalization for salvage therapy among adults with Philadelphia (Ph)-negative B-cell relapsed/refractory (R/R) acute lymphoblastic leukemia (ALL) in Spain. 19th ISPOR Annual European Congress 2016.
42. Barlev A, Lin VW, Song X. Burden of hospitalization in relapsed acute lymphoblastic leukemia. Curr Med Res Opin 2016;32:1209-12.
43. Lazarus HM, Advani AS. When, how, and what cell source for hematopoietic cell transplantation in first complete remission adult acute lymphoblastic leukemia? Hematology Am Soc Hematol Educ Program 2012;2012:382-8.
44. Fielding AK. Current treatment of Philadelphia chromosome-positive acute lymphoblastic leukemia. Hematology Am Soc Hematol Educ Program 2011;2011:231-7.
45. Cancer Research UK. UKALL14: a randomized trial for adults with newly diagnosed acute lymphoblastic leukaemia: trial protocol, 2012.
46. Gokbuget N, Hoelzer D. Treatment of adult acute lymphoblastic leukemia. Semin Hematol 2009;46:64-75.
47. Hummel H-D, Topp M, Chang E, et al. Adverse events in adults with relapsed or refractory acute lymphoblastic leukemia (ALL): a literature review of recent clinical trials. Journal of Leukemia 2016;4.
48. Spinelli O, Peruta B, Tosi M, et al. Clearance of minimal residual disease after allogeneic stem cell transplantation and the prediction of the clinical outcome of adult patients with high-risk acute lymphoblastic leukemia. Haematologica 2007;92:612-8.
49. Daikeler T, Tichelli A, Passweg J. Complications of autologous hematopoietic stem cell transplantation for patients with autoimmune diseases. Pediatr Res 2012;71:439-444.
50. Nagafuji K, Miyamoto T, Eto T, et al. Monitoring of minimal residual disease (MRD) is useful to predict prognosis of adult patients with Ph-negative ALL: results of a prospective study (ALL MRD2002 Study). J Hematol Oncol 2013;6:14.
51. National Institute for Health and Care Excellence. Pegaspargase for treating acute lymphoblastic leukaemia [TA408]. 2016.
52. Amgen I. Data on File: Minimal residual disease testing patterns and healthcare utilization for adults with B-precursor acute lymphoblastic leukemia - results of a physician survey in the UK. 2017.
53. Hoelzer D. Personalized medicine in adult acute lymphoblastic leukemia. Haematologica 2015;100:855-8.
54. Bruggemann M, Schrauder A, Raff T, et al. Standardized MRD quantification in European ALL trials: proceedings of the Second International Symposium on MRD assessment in Kiel, Germany, 18-20 September 2008. Leukemia 2010;24:521-35.
55. Bassan R, Spinelli O. Minimal Residual Disease Monitoring in Adult ALL to Determine Therapy. Curr Hematol Malig Rep 2015;10:86-95.
56. Amgen Inc. Clinical Study Report: MT103-203 Addendum 1 (Key Secondary Analysis), 2016.
57. Amgen Inc. Clinical Study Report: MT103-202. 2013.
58. Amgen Inc. Observational Research Study Report: A Retrospective Analysis of Hematological Relapse Free Survival (RFS) and Overall Survival (OS) in Adult Patients with Philadelphia-Negative B-Precursor Acute Lymphoblastic Leukemia (ALL) in Complete Hematological Remission with Minimal Residual Disease (MRD). 2017.
59. Amgen Inc. Study Protocol: A confirmatory multicenter, single-arm study to assess the efficacy, safety, and tolerability of the BiTE antibody blinatumomab in adult patients with minimal residual disease (MRD) of B-precursor acute lymphoblastic leukemia. 2010.
60. Rubin DB. Inference and missing data. Biometrika 1976;63:581-590.
61. Gökbuget N, Dombret, H., Bonifacio, M., Reichle, A., Graux, C., Faul, C., Diedrich, H., Topp, M. S., Brüggemann, M., Horst, H. A., Stieglmaier, J., Wessels, H., Haddad, V., Zugmaier, G., Nagorsen, D., & Bargou, R. C. Long-Term Outcomes after Blinatumomab Treatment: Follow-up of a Phase 2 Study in Patients (Pts) with Minimal Residual Disease (MRD) Positive B-Cell Precursor Acute Lymphoblastic Leukemia (ALL). Blood 2015;126:680.
62. Fayers PM, Aaronson NK, Bjordal K, et al. The EORTC QLQ-C30 Scoring Manual (3rd Edition). 2001.

63. Hong F, Bosco JL, Bush N, et al. Patient self-appraisal of change and minimal clinically important difference on the European organization for the research and treatment of cancer quality of life questionnaire core 30 before and during cancer therapy. BMC Cancer 2013;13:165.
64. Amgen Inc. TOWER NICE Submission. 2016.
65. Amgen Inc. Data on File: Propensity Score Analysis of Relapse Free Survival and Overall Survival among Adult Patients with Minimal Residual Disease (MRD) of B-precursor Acute Lymphoblastic Leukemia. 2016.
66. National Institute for Health and Care Excellence. NICE DSU Technical Support Document 17. 2015.
67. Amgen Inc. Clinical Study Report: MT103-206. .
68. Amgen Inc. Clinical Study Report: MT103-211.
69. Amgen Inc. Clinical Study Report: TOWER.
70. Amgen Inc. Clinical Overview: Blinatumomab. Regulatory Submission Dossier.
71. Bader P, Kreyenberg H, Henze GH, et al. Prognostic value of minimal residual disease quantification before allogeneic stem-cell transplantation in relapsed childhood acute lymphoblastic leukemia: the ALL-REZ BFM Study Group. J Clin Oncol 2009;27:377-84.
72. Amgen Inc. Amgen Receives FDA Breakthrough Therapy Designation For Investigational BiTE® Antibody Blinatumomab In Acute Lymphoblastic Leukemia, 2016.
73. Food and Drug Administration. Highlights of Prescribing Information: BLINCYTO (blinautmomab) for injection, for intravenous use, 2014.
74. National Institute for Health and Care Excellence. Trametinib in combination with dabrafenib for treating unresectable or metastatic melanoma [TA396]. 2016.
75. National Institute for Health and Care Excellence. Blinatumomab for previously treated Philadelphia-chromosome-negative acute lymphoblastic leukaemia [TA450]. 2017.
76. Woods B, Sideris E, Palmer S, et al. NICE DSU technical support document 19: partitioned survival analysis for decision modelling in health care: a critical review. Report by the Decision Support Unit. 2 June 2017. Available at: http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2017/06/Partitioned-Survival-Analysis-final-report.pdf. Accessed: 12 june 2017.
77. Garcia-Albeniz X, Maurel J, Hernan MA. Why post-progression survival and post-relapse survival are not appropriate measures of efficacy in cancer randomized clinical trials. Int J Cancer 2015;136:2444-7.
78. Martin PJ, Counts GW, Jr., Appelbaum FR, et al. Life expectancy in patients surviving more than 5 years after hematopoietic cell transplantation. J Clin Oncol 2010;28:1011-6.
79. Wingard JR, Majhail NS, Brazauskas R, et al. Long-term survival and late deaths after allogeneic hematopoietic cell transplantation. J Clin Oncol 2011;29:2230-9.
80. Bhatia S, Francisco L, Carter A, et al. Late mortality after allogeneic hematopoietic cell transplantation and functional status of long-term survivors: report from the Bone Marrow Transplant Survivor Study. Blood 2007;110:3784-92.
81. Bhatia S, Robison LL, Francisco L, et al. Late mortality in survivors of autologous hematopoietic-cell transplantation: report from the Bone Marrow Transplant Survivor Study. Blood 2005;105:4215-22.
82. Socie G, Stone JV, Wingard JR, 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;341:14-21.
83. Pond GR, Lipton JH, Messner HA. Long-term survival after blood and marrow transplantation: comparison with an age- and gender-matched normative population. Biol Blood Marrow Transplant 2006;12:422-9.
84. National Institute for Health and Care Excellence (NICE). Appraisal consultation document for inotuzumab ozogamicin for treating relapsed or refractory B-cell acute lymphoblastic leukaemia. Available at: https://www.nice.org.uk/guidance/gid-ta10091/documents/appraisal-consultation-document. Accessed: 17 July 2017.
85. National Institute for Health and Care Excellence (NICE). Single Technology Appraisal. Inotuzumab ozogamicin for treating relapsed or refractory B-cell acute lymphoblastic leukaemia [ID893]. Committee Papers. Available at: https://www.nice.org.uk/guidance/gid-ta10091/documents/committee-papers-2. Accessed: 17 July 2017.

86. National institute for Health and Care Excellence. Guide to the methods of technology appraisal. 2013.
87. Jackson C. Package 'flexsurv'. Flexible Parametric Survival and Multi-State Models. May 11, 2016. Available at: https://cran.r-project.org/web/packages/flexsurv/flexsurv.pdf. Accessed: September 26, 2016.
88. Lambert P. Modeling of the cure fraction in survival studies. The Stata Journal 2007;7:1-25.
89. Othus M, Bansal A, Koepl L, et al. Accounting for Cured Patients in Cost-Effectiveness Analysis. Value in Health 2017;20:705-709.
90. Bagust A, Beale S. Survival analysis and extrapolation modeling of time-to-event clinical trial data for economic evaluation: an alternative approach. Med Decis Making 2014;34:343-51.
91. Latimer NR. Survival analysis for economic evaluations alongside clinical trials--extrapolation with patient-level data: inconsistencies, limitations, and a practical guide. Med Decis Making 2013;33:743-54.
92. Raftery AE. Bayesian Model Selection in Social Research. Sociol Methodol 1995;25:111-1163.
93. Longworth L, Rowen D. Mapping to obtain EQ-5D utility values for use in NICE health technology assessments. Value Health 2013;16:202-10.
94. Pidala J, Anasetti C, Jim H. Health-related quality of life following haematopoietic cell transplantation: patient education, evaluation and intervention. Br J Haematol 2010;148:373-85.
95. Pidala J, Anasetti C, Jim H. Quality of life after allogeneic hematopoietic cell transplantation. Blood 2009;114:7-19.
96. Korol EE, Wang S, Johnston K, et al. Health-Related Quality of Life of Patients with Acute Myeloid Leukemia: A Systematic Literature Review. Oncol Ther 2017;5:1-16.
97. Kurosawa S, Yamaguchi T, Mori T, et al. Patient-reported quality of life after allogeneic hematopoietic cell transplantation or chemotherapy for acute leukemia. Bone Marrow Transplant 2015;50:1241-9.
98. Leunis A, Redekop WK, Uyl-de Groot CA, et al. Impaired health-related quality of life in acute myeloid leukemia survivors: a single-center study. Eur J Haematol 2014;93:198-206.
99. Slovacek L, Slovakova B, Jebavy L. [Health and psychosocial aspects of quality of life in patients undergoing autologous progenitor stem cell transplantation]. Cas Lek Cesk 2006;145:949-52; discussion 953-4.
100. Kurosawa S, Yamaguchi H, Yamaguchi T, et al. Decision Analysis of Postremission Therapy in Cytogenetically Intermediate-Risk Acute Myeloid Leukemia: The Impact of FLT3 Internal Tandem Duplication, Nucleophosmin, and CCAAT/Enhancer Binding Protein Alpha. Biol Blood Marrow Transplant 2016;22:1125-32.
101. Xu Y, Hijazi Y, Wolf A, et al. Physiologically Based Pharmacokinetic Model to Assess the Influence of Blinatumomab-Mediated Cytokine Elevations on Cytochrome P450 Enzyme Activity. CPT Pharmacometrics Syst Pharmacol 2015;4:507-15.
102. Kohnke T, Krupka C, Tischer J, et al. Increase of PD-L1 expressing B-precursor ALL cells in a patient resistant to the CD19/CD3-bispecific T cell engager antibody blinatumomab. J Hematol Oncol 2015;8:111.
103. MedicinesComplete (BNF): Royal Pharmaceutical Society, 2017.
104. Gokbuget N, Kelsh M, Chia V, et al. Blinatumomab vs historical standard therapy of adult relapsed/refractory acute lymphoblastic leukemia. Blood Cancer J 2016;6:e473.
105. Unrelated Donor Stem Cell Transplantation in the UK. In: Committee USCSO, ed: NHS Blood and Transplant, 2014.
106. Curtis L, Burns A. Unit Costs of Health & Social Care 2016. In: Unit PSSR, ed: University of Kent, 2016.
107. National Institute for Health and Care Excellence. Final appraisal determination – Blinatumomab for treating Philadelphia-chromosome-negative relapsed or refractory acute lymphoblastic leukaemia Issue date: April 2017. Available at: https://www.nice.org.uk/guidance/gid-ta10093/documents/final-appraisal-determination-document. Accessed: June 18, 2017.

108. Addicott R, Dewar S. Improving choice at end of life: a descriptive analysis of the impact and costs of the Marie Cure Delivering Choice Programme in Lincolnshire. London, UK: King's Fund, 2008.
109. Understanding the cost of end of life care in different settings: Marie Curie Cancer Care, 2012.
110. Tinkham A, Kaner C. Exploring exploratory testing. Paper presented atthe STAR East 2003 Conference, Orlando, FL, USA, 2003. http://www.testingeducation.org/articles/exploring_exploratory_testing_star_east_2003_paper.pdf). Accessed 2 August 2017.
111. Dasbach EJ, Elbasha EH. Verification of Decision-Analytic Models for Health Economic Evaluations: An Overview. Pharmacoeconomics 2017;35:673-683.
112. Tappenden P, Chilcott JB. Avoiding and identifying errors and other threats to the credibility of health economic models. Pharmacoeconomics 2014;32:967-79.
113. Moher D, Liberati A, Tetzlaff J, et al. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLOS Medicine 2009;6:e1000097.
114. The Cochrane Collaboration. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0, 2011.
115. Amgen Inc. Data on File: Clinical SLR Protocol. 2017.
116. Gokbuget N, Kneba M, Raff T, et al. Adult patients with acute lymphoblastic leukemia and molecular failure display a poor prognosis and are candidates for stem cell transplantation and targeted therapies. Blood 2012;120:1868-76.

Appendices

Appendix C Summary of product characteristics (SmPC) and European public assessment report (EPAR)

The SmPC for blinatumomab is included in the reference pack supplied with this submission.[8]

The EPAR is unavailable at the time of this submission; a European marketing authorisation application for blinatumomab for MRD+ BCP-ALL was submitted in March 2017, and it is anticipated that the Committee for Medicinal Products for Human Use (CHMP) will adopt a positive opinion for this MAA in January 2018 for the indication of adults with MRD+ B-precursor ALL. The EPAR will therefore be available after the marketing authorisation is granted.

Appendix D Identification, selection and synthesis of clinical evidence

Systematic Literature Review

A comprehensive systematic literature review (SLR) was conducted on 19th May 2017 to identify RCTs and observational studies reporting the efficacy and safety of current treatments for adult patients with MRD+ BCP-ALL. The SLR was conducted in accordance with the requirements of NICE, Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines and The Cochrane Collaboration.[86, 113, 114]

Search Strategies

The following databases were searched for published literature on efficacy, safety, humanistic burden, and economic outcomes for adults with MRD-positivity after treatment:

- MEDLINE In-Process via PubMed
- Embase
- The Cochrane Library
- Central Register of Controlled Trials (CENTRAL)
- Cochrane Database of Systematic Reviews (CDSR)
- Database of Abstracts of Reviews of Effectiveness (DARE)
- NHS Economic Evaluation Database (NHS EED)
- Health Technology Appraisals (HTA)

These databases were searched via the EMBASE®, PubMed, Cochrane Library and York Centre for Reviews and Dissemination platforms by developing search strategies combining free-text search terms and controlled vocabulary terms, which are presented in their entirety in Table 68, Table 69, Table 70 and Table 71, respectively. These strategies were also used to identify relevant studies for the cost-effectiveness, HRQoL, and cost and healthcare resource use SLRs.

Table 68. PubMed Search Algorithm

#	Search Terms	Yields
1	"Precursor Cell Lymphoblastic Leukemia-Lymphoma"[Mesh] OR "acute lymphoblastic leukemia" OR "acute lymphoblastic leukaemia" OR "acute lymphocytic leukemia" OR "acute lymphocytic leukaemia" OR ((lymphocyt*[TIAB] OR lymphoblast*[TIAB] OR lymphat*[TIAB] OR lymphoid*[TIAB]) AND (leukemi*[TIAB] OR leukaemi*[TIAB]) AND acute[TIAB])	45,725
2	"minimal residual disease" OR "minimal residual disease positive" OR MRD OR "residual malignant cells" OR "low-level disease"	6,515
3	#1 AND #2	1,761
4	pediatric[tiab] OR paediatric[tiab] OR children[tiab] OR childhood[tiab] OR child[tiab] OR adolescent[tiab] OR toddler[tiab] OR infant[tiab] OR newborn[tiab] OR kid[tiab]	1,543,708
5	#3 NOT #4	830

6	#5 limit to humans	711
7	#6 limit to articles published in English	631
8	"case reports"[ptyp] OR review[ptyp] OR editorial[ptyp] OR news[ptyp] OR letter[ptyp]	3,808,752
9	#7 NOT #8	403

Source: Clinical SLR Protocol[115]

Table 69. Embase Search Algorithm

#	Search Terms	Yields
1	'acute lymphoblastic leukemia'/exp OR 'acute lymphoblastic leukemia' OR 'acute lymphocytic leukemia'/exp OR 'acute lymphocytic leukemia'	62,646
2	((lymphocyt* OR lymphoblast* OR lymphat* OR lymphoid*) NEAR/1 (leukemi* OR leukaemi*)):ab,ti AND (acute NEAR/3 (lymphocyt* OR lymphoblast* OR lymphat* OR lymphoid*)):ab,ti	44,742
3	#1 OR #2	66,763
4	'minimal residual disease' OR 'minimal residual disease positive' OR MRD OR 'residual malignant cells' OR 'low-level disease'	26,509
5	#3 AND #4	3,711
6	pediatric:ab,ti OR paediatric:ab,ti OR children:ab,ti OR childhood:ab,ti OR child:ab,ti OR adolescent:ab,ti OR toddler:ab,ti OR infant:ab,ti OR newborn:ab,ti OR kid:ab,ti	1,917,455
7	#5 NOT #6	1,770
8	#7 limit to humans	1,667
9	#8 limit to articles published in English	1,546
10	#9 limited to articles and articles in press	632
11	#10 NOT 'case report'/de	551

Source: Clinical SLR Protocol[115]

Table 70. Cochrane Library Search Algorithm

#	Search Terms	Yields
1	MeSH descriptor: [Precursor Cell Lymphoblastic Leukemia-Lymphoma] explode all trees	857
2	"acute lymphoblastic leukemia" or "acute lymphoblastic leukaemia" or "acute lymphocytic leukemia" or "acute lymphocytic leukaemia" or ((lymphocyt* or lymphoblast* or lymphat* or lymphoid*) and (leukemi* or leukaemi*) and acute) [Search all fields; word variations have been searched]	2,756
3	#1 OR #2	2,819
4	"minimal residual disease" OR "minimal residual disease positive" OR MRD OR "residual malignant cells" OR "low-level disease" [Search all fields; word variations have been searched]	882
5	#3 and #4	172
6	pediatric or paediatric or children or childhood or child or adolescent or toddler or infant or newborn or kid:ti,ab,kw [Word variations have been searched]	199,469
7	#5 not #6	50

Source: Clinical SLR Protocol[115]

Table 71. York Centre for Reviews and Dissemination Search Algorithm

#	Search Terms	Yields
1	"acute lymphoblastic leukemia" [Any field] AND (cost or economic) [Any field]	14

In addition, unpublished "grey literature", i.e. publicly available data that are not published in a peer-reviewed journal, was also searched. Studies must have been published in English, but there were no geographic or temporal limits on inclusion. Searches of conference proceedings, limited to conferences from the past three years or three editions, were performed for the following:

- ASBMT – American Society for Blood and Marrow Transplantation
- ASCO – American Society of Clinical Oncology
- ASH – American Society of Hematology
- ECCO/ESMO – European Cancer Organisation/European Society for Medical Oncology
- EHA – European Hematology Association
- ISPOR – International Society for Pharmacoeconomics and Outcomes Research
- Additionally, the search was supplemented with unpublished Amgen studies considered relevant to the decision problem

Study Selection

Articles identified through the electronic database searches and hand searches were screened using a two-level selection and evaluation process. In the first level of review, the pre-defined inclusion and exclusion criteria were used to evaluate the titles/abstracts of records identified from the searches. Full-text articles were then retrieved and reviewed for abstracts that were deemed relevant during the first level of review. During the review process, records were screened by two independent reviewers and a third, senior reviewer reconciled any discrepancies between the screening results. Table 72 summarises the eligibility criteria used in the study selection process for the clinical studies.

Table 72. Eligibility criteria used in the study selection process across the clinical efficacy/safety SLR

	Inclusion Criteria	Exclusion Criteria
Population	Adult ALL patients with MRD positivity[i] after treatment[ii]	• Paediatric patients • MRD- ALL patients
Intervention/Comparator	Any interventional therapies	None
Outcomes	Clinical effectiveness and safety[iii] • OS • RFS • Event-free survival • MRD complete response rate • Duration of MRD response • Duration of haematologic response • Rate of transplant • Mortality following transplant	Non-clinical outcomes, such as those in pharmacodynamics or *in vitro* studies

	• Treatment-related mortality • Serious adverse events • Grade 3 or 4 adverse events (list to be determined based on the most commonly reported) • Discontinuations due to adverse events • PROs	
Study Design	• RCTs of at least 10 patients per arm • Single-arm clinical trials of at least 10 patients • Prospective and retrospective observational studies of at least 10 patients	• Case studies and studies evaluating fewer than 10 patients • Letters, narrative reviews, expert opinions, etc.

Footnotes: [i]Definition of MRD was captured as reported by each study author. It is noted that a number of different technologies are available to measure MRD, with varying specificity and sensitivity. The most common methods include: multicolour flow cytometry to detect abnormal immunophenotypes, with a sensitivity of 10^{-3} to 10^{-4} for 3 to 4 colour flow cytometry and 10^{-4} to 10^{-5} for 6 to 9 color flow cytometry; RT-qPCR assays to detect clonal rearrangements in immunoglobulin (Ig) heavy chain genes, and/or TCR genes, with a sensitivity of 10^{-4} to 10^{-5}; RT-qPCR assays to detect fusion genes (e.g., BCR–ABL), with a sensitivity of 10^{-4} to 10^{-6}. [ii]For inclusion, studies must have involved adult ALL patients who are MRD+ after treatment; data for MRD- populations was extracted from any comparative studies that presented data for both groups. [iii]Observational studies only. [iii]Observational and interventional studies.
Abbreviations: SLR: systematic literature review; ALL: acute lymphoblastic leukaemia; MRD: minimal residual disease; HRQoL: health-related quality of life; PRO: patient-reported outcome; OS: overall survival; RFS: relapse-free survival; RCT: randomised clinical trial.

Other limits and considerations applied during study selection and review are provided in Table 73.

Table 73. Other limits/considerations used in the study selection process across the clinical efficacy/safety SLR

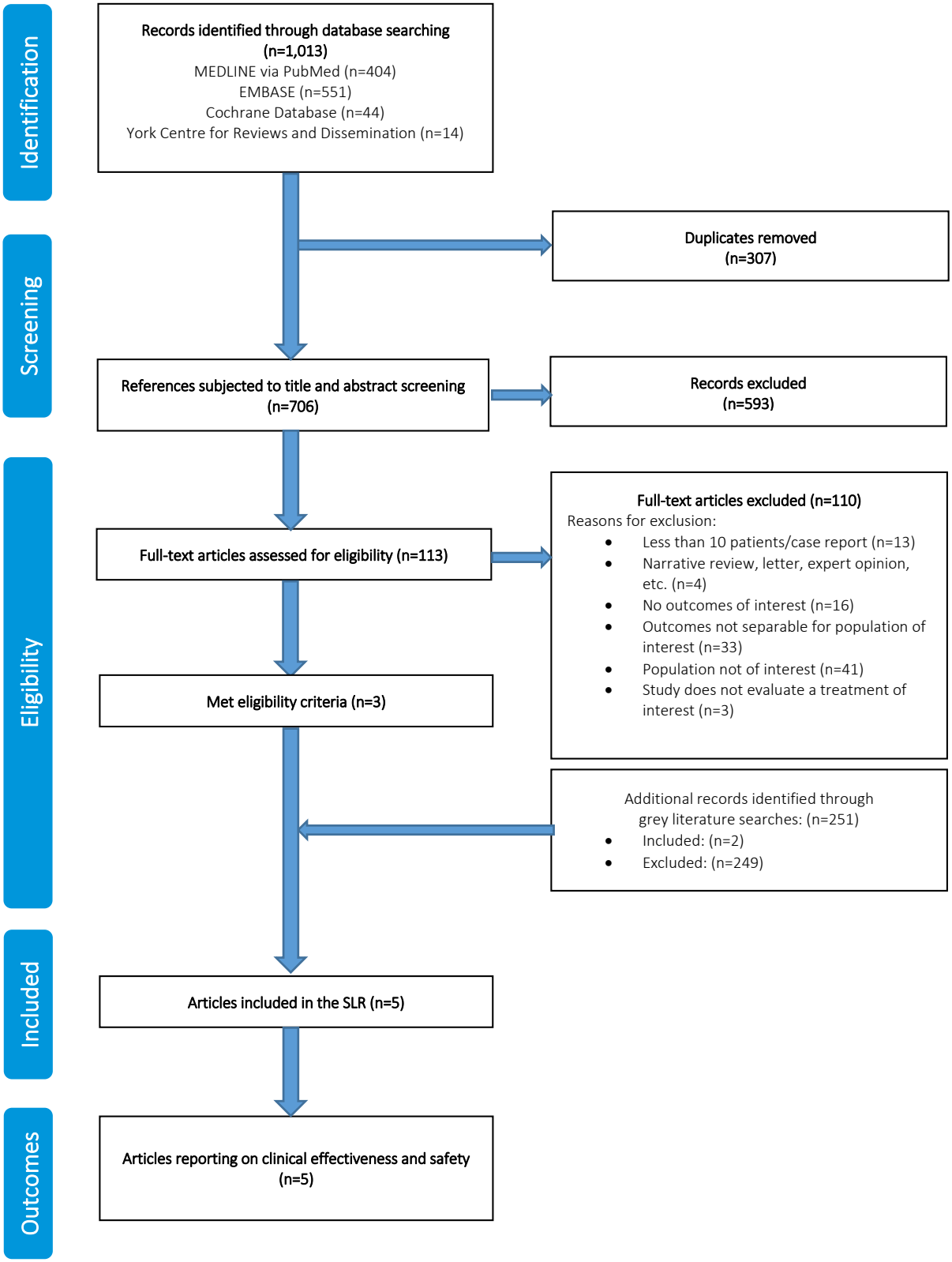
	Limits/considerations
Subgroups of interest	Data for the following subgroups were of interest and were extracted if reported in studies included in the SLR: • CR1 ALL patients who achieved haematological complete remission after front line treatment and are MRD+ • CR2+ ALL patients who achieved haematological complete remission after salvage treatment and are MRD+
Timeframe	The following temporal limits were applied: • Literature databases & registries: No date restrictions • Grey literature & scientific meetings: Past 3 editions or past 3 years
Setting	Articles in English were included with no limits applied to the geographic location or setting in which the study was conducted. As blinatumomab is the only treatment within this indication, no other studies were available that would not have been captured within this study setting.

Abbreviations: SLR: systematic literature review; CR: complete remission; MRD: minimal residual disease.

Flow diagram for clinical studies

The following PRISMA diagram, presented in Figure 46, shows the SLR process including the total number of records identified in the searches, and the reasons for study exclusion for clinical studies. A complete list of excluded publications is provided in Appendix M.

Figure 46. PRISMA diagram for clinical SLR



Abbreviations: PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; SLR: systematic literature review; ALL: acute lymphoblastic leukaemia; MRD: minimal residual disease.

The SLR identified 3 relevant studies, which are discussed in more detail in Section B.2.2. PRISMA diagrams for the cost-effectiveness, HRQoL, and cost and healthcare resource use SLRs are included in Appendix G, Appendix H, and Appendix I, respectively.

Quality Assessment

Full quality assessments of the 3 relevant studies are included below in Table 74.

Table 74. ROBINS-I Risk of Bias Assessment

Signalling questions	Response options	Response		
		Pilot study (MT103-202)	BLAST (MT103-203)	Historical Comparator (20120148)
BIAS DUE TO CONFOUNDING				
1.1 Is there potential for confounding of the effect of intervention in this study? **If N/PN to 1.1: the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered**	Y / PY / PN / N	N – the study evaluates a single treatment	N – the study evaluates a single treatment	N – the study does not compare treatments
If Y/PY to 1.1: determine whether there is a need to assess time-varying confounding:				
1.2. Was the analysis based on splitting participants' follow up time according to intervention received? **If N/PN, answer questions relating to baseline confounding (1.4 to 1.6)** **If Y/PY, go to question 1.3.**	NA / Y / PY / PN / N / NI	NA	NA	NA
1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome? **If N/PN, answer questions relating to baseline confounding (1.4 to 1.6)** **If Y/PY, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8)**	NA / Y / PY / PN / N / NI	NA	NA	NA
Questions relating to baseline confounding only:				
1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains?	NA / Y / PY / PN / N / NI	NA	NA	NA

1.5. If Y/PY to 1.4: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?	NA / Y / PY / PN / N / NI	NA	NA	NA
1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention?	NA / Y / PY / PN / N / NI	NA	NA	NA
Questions relating to baseline and time-varying confounding:				
1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?	NA / Y / PY / PN / N / NI	NA	NA	NA
1.8. If Y/PY to 1.7: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?	NA / Y / PY / PN / N / NI	NA	NA	NA
RISK OF BIAS JUDGEMENT	LOW / MODERATE / SERIOUS / CRITICAL / NI	LOW	LOW	LOW
BIAS IN SELECTION OF PARTICIPANTS INTO THE STUDY				
2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention?	Y / PY / PN / N / NI	N- patients were selected for inclusion prior to receiving study treatment	N- patients were selected for inclusion prior to receiving study treatment	Yes – patients were selected from retrospective cohorts.
If N/PN to 2.1: go to 2.4				
2.2. If Y/PY to 2.1: Were the post-intervention variables that influenced selection likely to be associated with intervention?	NA / Y / PY / PN / N / NI	NA	NA	N
2.3 If Y/PY to 2.2: Were the post-intervention variables that influenced	NA / Y / PY / PN / N / NI	NA	NA	NA

selection likely to be influenced by the outcome or a cause of the outcome?				
2.4. Do start of follow-up and start of intervention coincide for most participants?	Y / PY / PN / N / NI	Y	Y	NA
2.5. If Y/PY to 2.2 and 2.3, or N/PN to 2.4: Were adjustment techniques used that are likely to correct for the presence of selection biases?	NA / Y / PY / PN / N / NI	NA	NA	NA
RISK OF BIAS JUDGEMENT	LOW / MODERATE / SERIOUS / CRITICAL / NI	LOW	LOW	LOW
BIAS IN CLASSIFICATION OF INTERVENTIONS				
3.1 Were intervention groups clearly defined?	Y / PY / PN / N / NI	Y	Y	NA (retrospective non-interventional cohort study)
3.2 Was the information used to define intervention groups recorded at the start of the intervention?	Y / PY / PN / N / NI	Y	Y	NA
3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?	Y / PY / PN / N / NI	N	N	NA
RISK OF BIAS JUDGEMENT	LOW / MODERATE / SERIOUS / CRITICAL / NI	LOW	LOW	LOW
BIAS DUE TO DEVIATIONS FROM INTENDED INTERVENTIONS				
If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2				
4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice?	Y / PY / PN / N / NI	N	N	NA

4.2. If Y/PY to 4.1: Were these deviations from intended intervention unbalanced between groups and likely to have affected the outcome?	NA / Y / PY / PN / N / NI	NA	NA	NA
If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6				
4.3. Were important co-interventions balanced across intervention groups?	Y / PY / PN / N / NI	NA	NA	NA
4.4. Was the intervention implemented successfully for most participants?	Y / PY / PN / N / NI	NA	NA	NA
4.5. Did study participants adhere to the assigned intervention regimen?	Y / PY / PN / N / NI	NA	NA	NA
4.6. If N/PN to 4.3, 4.4 or 4.5: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?	NA / Y / PY / PN / N / NI	NA	NA	NA
RISK OF BIAS JUDGEMENT	LOW / MODERATE / SERIOUS / CRITICAL / NI	NI	NI	NI
BIAS DUE TO MISSING DATA				
5.1 Were outcome data available for all, or nearly all, participants?	Y / PY / PN / N / NI	Y	Y	Y
5.2 Were participants excluded due to missing data on intervention status?	Y / PY / PN / N / NI	N	N	Y (missing baseline MRD status for some patients)
5.3 Were participants excluded due to missing data on other variables needed for the analysis?	Y / PY / PN / N / NI	N	N	N
5.4 If PN/N to 5.1, or Y/PY to 5.2 or 5.3: Are the proportion of participants and reasons for missing data similar across interventions?	Y / PY / PN / N / NI	NA	NA	NA

5.5 If PN/N to 5.1, or Y/PY to 5.2 or 5.3: Is there evidence that results were robust to the presence of missing data?	Y / PY / PN / N / NI	NA	NA	NA
RISK OF BIAS JUDGEMENT	LOW / MODERATE / SERIOUS / CRITICAL / NI	LOW	LOW	LOW
BIAS IN MEASUREMENT OF OUTCOMES				
6.1 Could the outcome measure have been influenced by knowledge of the intervention received?	Y / PY / PN / N / NI	N	N	N
6.2 Were outcome assessors aware of the intervention received by study participants?	Y / PY / PN / N / NI	Y	Y	Y
6.3 Were the methods of outcome assessment comparable across intervention groups?	Y / PY / PN / N / NI	NA	NA	N
6.4 Were any systematic errors in measurement of the outcome related to intervention received?	Y / PY / PN / N / NI	NA	NA	N
RISK OF BIAS JUDGEMENT	LOW / MODERATE / SERIOUS / CRITICAL / NI	LOW	LOW	LOW
BIAS IN SELECTION OF THE REPORTED RESULT				
Is the reported effect estimate likely to be selected, on the basis of the results, from...				
7.1. ... multiple outcome measurements within the outcome domain?	Y / PY / PN / N / NI	N	N	N
7.2 ... multiple analyses of the intervention-outcome relationship?	Y / PY / PN / N / NI	N	N	N

7.3 ... different subgroups?	Y / PY / PN / N / NI	N	N	N
RISK OF BIAS JUDGEMENT	LOW / MODERATE / SERIOUS / CRITICAL / NI	LOW	LOW	LOW
OVERALL RISK OF BIAS JUDGEMENT	**LOW / MODERATE / SERIOUS / CRITICAL / NI**	**LOW**	**LOW**	**LOW**

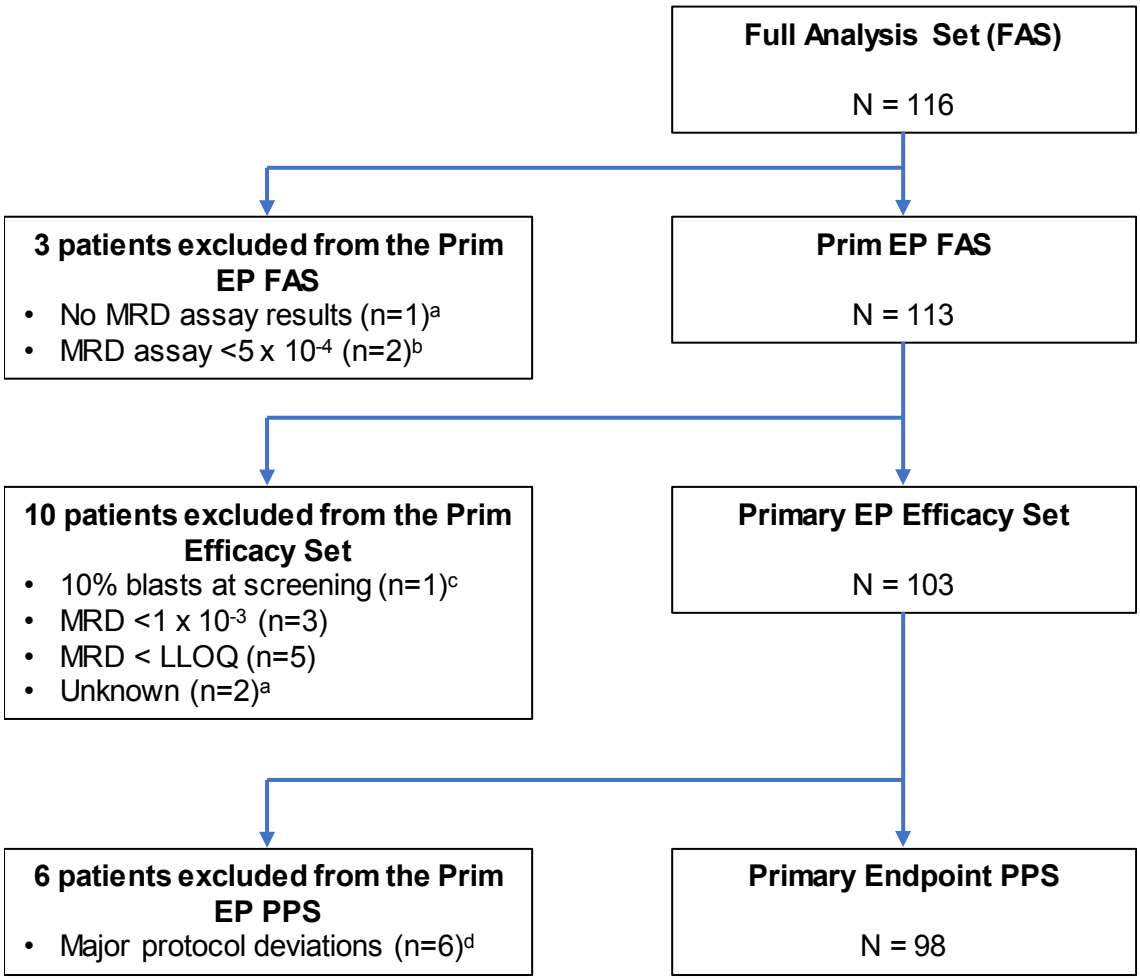
Abbreviations: ROBINS-I: Risk Of Bias In Non-randomized Studies – of Interventions.

Participant flow

Diagrams of the primary endpoint analysis sets and key secondary endpoint analysis sets are presented in Figure 47 and Figure 48, respectively. In total, 211 patients were enrolled (i.e. screened), of which 116 patients received at least 1 infusion of blinatumomab and were therefore included in the FAS. Overall, 113 patients (97.4%; 113/116) were included in the primary analysis, while 74 patients (63.8%; 74/116) were included in the HSCT Secondary Efficacy Full Analysis Set (HSCT Sec EP FAS).

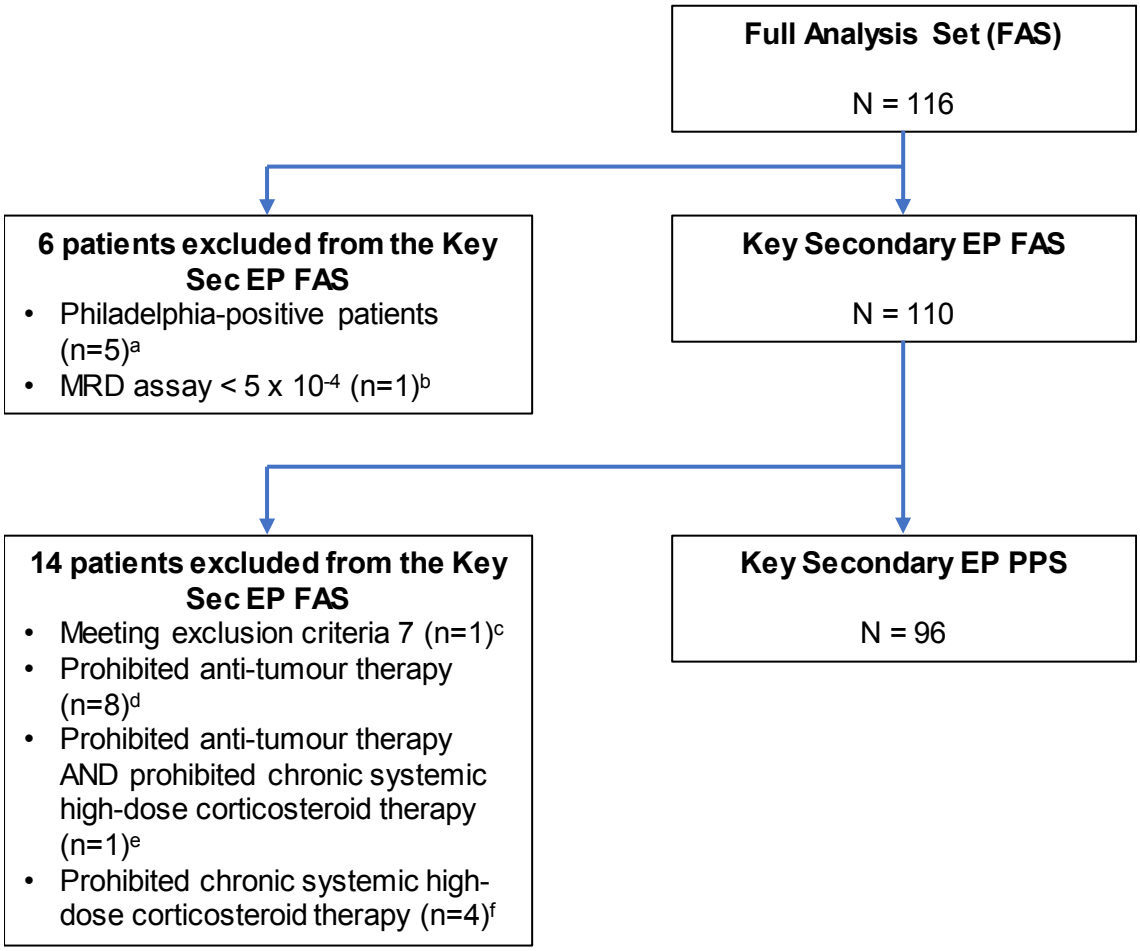
Eighty-three patients (71.6%; 83/116) completed the core study. Reasons for not completing the core study included adverse events (17.2%; 20/116), disease relapse (8.6%; 10/116), physician decision (1.7%; 2/116), and "other" reason (0.9%, 1/116). Of the patients who ended the core study, the median duration of the core study was 2.7 months (range: 0 to 7 months). At the time of the data cut-off, 53.4% (62/116) of patients are continuing the study (i.e. participating in the survival follow-up) and 46.6% (54/116) of patients ended the study. Of the patients who ended the study, 45.7% (53/116) had died, and 0.9% (1/116) withdrew. The median total time on the study for all 116 patients was 18.3 months (range: 1 to 54 months).

Figure 47. BLAST primary endpoint analysis sets



Footnotes: [a]Patient 1311-004, [b]Patient 1301-002 and Patient 1407-005, [c]Patient 1018-002, [d]Patient 1303-005 was already excluded from the Prim Efficacy FAS (MRD < LLOQ).
Abbreviations: Prim EP FAS: Primary efficacy endpoint FAS; MRD: Minimal Residual Disease; Prim EP PPS: Primary efficacy endpoint per protocol set.
Source: BLAST Key Secondary Analysis CSR Figure 9-1[56]

Figure 48. BLAST key secondary endpoint analysis sets



Footnotes: [a]Patients 1015-002, 1016-008, 1202-005 and 1501-002 (Patients 1303-005 and 1015-002 also met exclusion criteria 7, and Patient 1501-002 also had prohibited anti-tumour therapy and tyrosine kinase inhibitors); [b]Patient 1018-002; [c]Patient 1001-008; [d]Patients 1002-003, 1016-004, 1201-003, 1201-003, 1201-003, 1201-005, 1302-001, 1501-003, 2101-001, and 2106-001; [e]Patient 1210-002; [f]Patients 1002-001, 1012-007, 1022-004, and 1607-001.
Abbreviations: Key Secondary EP FAS: key secondary efficacy endpoint FAS; MRD: Minimal Residual Disease; Key Secondary EP PPS: key secondary efficacy endpoint per protocol set.
Source: BLAST Key Secondary Analysis CSR Figure 9-2[56]

Appendix E Subgroup analysis

Results of the prespecified subgroup analyses for baseline variables, as listed in Table 26, are provided in this section for the primary outcome (MRD CR within 1 cycle), key secondary outcome (haematological RFS rate), other secondary outcomes (OS, TTHR), and HSCT status.

Effect of baseline variables on complete MRD response within 1 cycle

None of the subgroups tested were determined to have a statistically significant effect on complete MRD response, demonstrating the robust treatment effect provided by blinatumomab, even across older patients and those with a high level of MRD; both populations which currently have limited treatment options.

Effect of baseline variables on RFS

Only relapse history was considered to have a statistically significant effect on RFS (P = 0.0044), as presented in Table 75. Patients who were in the first CR at the time of start of treatment with blinatumomab had a significantly longer median RFS (24.6 months) than patients who were in second or third CR (11.0 months). Additionally, the 18 months KM estimate was higher for patients in the first CR compared with patients in the second or third CR (62% versus 34%, respectively). Patients in the second or third CR who have not yet relapsed or died at a given time point had a more than 2-fold higher chance of having haematologic relapse or death at the next time point compared with patients in the first CR (HR: 2.09, 95% CI: 1.26, 3.48).

Table 75. Effect of relapse history on RFS

Relapse history	Events[a]/ patients	Median (months)	18 months KM estimate	Hazard ratio[b] (95% CI)	P-value
Patients in 1st CR	36/75	24.6	0.62	Reference	0.0044
Patients in 2nd and 3rd CR	26/35	11.0	0.34	2.09 (1.26, 3.48)	

Footnotes: [a]Events are relapses, secondary leukaemia and deaths; [b]The hazard and hazard ratio estimates are obtained from the Cox Proportional Hazard Model. A hazard ratio < 1.0 indicates a lower average event rate and a longer RFS compared to the reference group.
Abbreviations: RFS: Relapse-Free Survival; CI: Confidence Interval; CR: Complete Remission.
Source: BLAST Key Secondary Analysis CSR Table 14-4.2.4[56]

Effect of baseline variables on OS

Only the Philadelphia status of the disease (Ph− disease or Ph+ disease) was considered to have a statistically significant effect on OS (P = 0.017), as presented in Table 76. Patients with Ph− disease had a significantly longer median OS (36.5 months versus 7.2 months) and 18 months KM estimate (67% versus 20%) than Ph+ patients. Patients with Ph+ disease who had not yet died by a given time had more than 3 times the chance of death at the next time point compared with Ph− patients (HR: 3.51, 95% CI: 1.26, 9.83). However, only 5 Ph+ patients were enrolled in the study, compared with 111 Ph− patients, and all 5 Ph+ patients were in CR 2 or CR 3 rather than CR 1 therefore the interpretation of this subgroup analysis should be treated with caution.

Table 76. Effect of Philadelphia status on OS

Philadelphia status	Events[a]/ patients	Median (months)	18 months KM estimate	Hazard ratio[b] (95% CI)	P-value
Positive	4/5	7.2	0.20	3.51 (1.26, 9.83)	0.017
Negative	49/111	36.5	0.67	Reference	

Footnotes: [a]Events are deaths; [b]The hazard and hazard ratio estimates are obtained from the Cox Proportional Hazard Model. A hazard ratio < 1.0 indicates a lower average event rate and a longer OS compared to the reference group.
Abbreviations: OS: Overall survival; KM: Kaplan-Meier; CI: Confidence Interval.
Source: BLAST Key Secondary Analysis CSR Table 14-4.3.4[56]

Effect of baseline variables on TTHR

Only relapse history was considered to have a statistically significant effect on TTHR (P = 0.0031), as presented in Table 77. Patients who were in the first CR at the time of start of treatment with blinatumomab had a longer median time to relapse (not estimable versus 15.0 months) and 18 months KM estimate (76% versus 48%) than patients in second or third CR. Patients in the second or third CR who have not yet relapsed at a given time point were at a 2.6-fold higher chance of haematologic relapse at the next time point compared with patients in first CR (HR: 2.60, 95% CI: 1.38, 4.89).

Table 77. Effect of relapse history on TTHR

Relapse history	Events[a]/ patients	Median (months)	18 months KM estimate	Hazard ratio[b] (95% CI)	P-value
Patients in 1st CR	30/75	36.5	0.69	Reference	0.087
Patients in 2nd and 3rd CR	23/41	19.1	0.56	1.61 (0.93, 2.77)	

Footnotes: [a]Events are relapses and secondary leukaemia; [b]The hazard and hazard ratio estimates are obtained from the Cox Proportional Hazard Model. A hazard ratio < 1.0 indicates a lower average event rate and a longer time to relapse compared to the reference group.
Abbreviations: TTHR: time to haematological relapse; KM: Kaplan-Meier; CI: Confidence Interval; CR: Complete remission.
Source: BLAST Key Secondary Analysis CSR Table 14-4.4.4[56]

Effect of baseline variables on HSCT

Baseline disease characteristics, broken out by whether the patients received HSCT, are presented in Table 78. More patients who received HSCT before relapse or never relapsed (72.4% [55/76]) or who did not receive HSCT (61.5% [16/26]) were in the first CR compared with patients who received HSCT after relapse (28.6% [4/14]). Similarly, fewer patients who received HSCT before relapse or never relapsed (11.8% [9/76]) or did not receive HSCT (11.5% [3/26]) had a > 30,000/mm^3 WBC count at first diagnosis.

Table 78. Effect of baseline variables on HSCT

Baseline variables	Patients who received HSCT before relapse or never relapsed (N = 76)	Patients who received HSCT after relapse (N = 14)	Patients who did not receive HSCT (N = 26)

Relapse history			
1st CR	55 (72.4)	4 (28.6)	16 (61.5)
2nd CR	21 (27.6)	10 (71.4)	8 (30.8)
3rd CR	0 (0.0)	0 (0.0)	2 (7.7)
WBC at first diagnosis			
≤30,000/mm^3	51 (67.1)	6 (42.9)	21 (80.8)
>30,000/mm^3	9 (11.8)	6 (42.9)	3 (11.5)
Unknown	16 (21.1)	2 (14.3)	2 (7.7)

Abbreviations: HSCT: haematopoietic stem cell transplant; CR: Complete response; WBC: white blood cell.
Source: BLAST Key Secondary Analysis CSR Table 14-4.7.4[56]

Appendix F Adverse reactions

No additional data or studies reporting adverse reactions were identified.

Appendix G Published cost-effectiveness studies

A single systematic literature review was performed to identify the available evidence on treatment patterns, humanistic and economic burden, clinical and HRQoL outcomes, and economic evidence relating to the treatment of adults with MRD+ BCP-ALL. The methodology of this SLR is described in detail in Appendix D.

Study Selection

Study selection was performed as described in Appendix D. Table 79 summarises the eligibility criteria used in the study selection process for the economic studies.

Table 79. Eligibility criteria used in the study selection process in the economic SLR

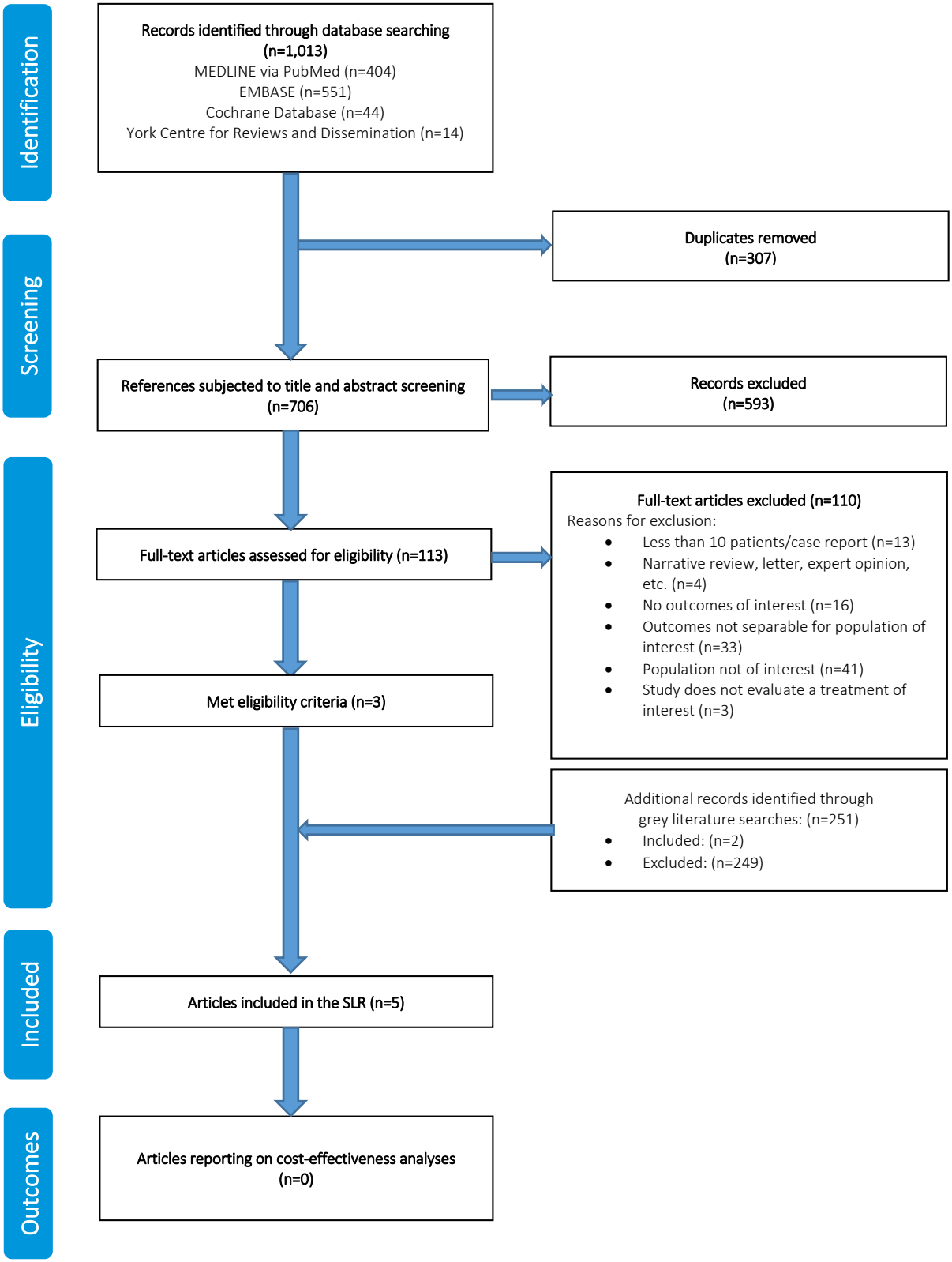
	Inclusion Criteria	Exclusion Criteria
Population	Adult ALL patients with MRD-positivity[i] after treatment[ii]	• Paediatric patients • MRD- ALL patients
Intervention/Comparator	Any interventional therapies	None
Outcomes	Cost effectiveness • Measures of cost effectiveness (e.g. costs per QALY)[iv]	Non-economic outcomes
Study Design	• Economic analyses and HTA reports	Non-economic study designs

Footnotes: [i]Definition of MRD was captured as reported by each study author. It is noted that a number of different technologies are available to measure MRD, with varying specificity and sensitivity. The most common methods include: multicolour flow cytometry to detect abnormal immunophenotypes, with a sensitivity of 10^{-3} to 10^{-4} for 3 to 4 colour flow cytometry and 10^{-4} to 10^{-5} for 6 to 9 color flow cytometry; RT-qPCR assays to detect clonal rearrangements in immunoglobulin (Ig) heavy chain genes, and/or TCR genes, with a sensitivity of 10^{-4} to 10^{-5}; RT-qPCR assays to detect fusion genes (e.g., BCR–ABL), with a sensitivity of 10^{-4} to 10^{-6}. [ii]For inclusion, studies must have involved adult ALL patients who are MRD+ after treatment; data for MRD- populations was extracted from any comparative studies that presented data for both groups. [iii]Observational studies only. [iv]Observational and interventional studies. [iv]Economic analyses.
Abbreviations: SLR: systematic literature review; ALL: acute lymphoblastic leukaemia; MRD: minimal residual disease; HCRU: healthcare resource utilisation; QALY: quality-adjusted life year; HTA: health technology assessment.

Flow diagram for cost-effectiveness studies

The following PRISMA diagram, presented in Figure 49, shows the SLR process including the total number of records identified in the searches, and the reasons for study exclusion for cost effectiveness studies. A complete list of excluded publications is provided in Appendix M. No studies providing cost effectiveness evidence were identified in the SLR.

Figure 49. PRISMA diagram for cost effectiveness studies



Abbreviations: PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; SLR: systematic literature review; ALL: acute lymphoblastic leukaemia; MRD: minimal residual disease.

Appendix H Health-related quality-of-life studies

A single systematic literature review was performed to identify the available evidence on treatment patterns, humanistic and economic burden, clinical and HRQoL outcomes, and economic evidence relating to the treatment of adults with MRD+ BCP-ALL. The methodology of this SLR is described in detail in Appendix D.

Study Selection

Study selection was performed as described in Appendix D. Table 80 summarises the eligibility criteria used in the study selection process for the HRQoL studies.

Table 80. Eligibility criteria used in the study selection process in the HRQoL SLR

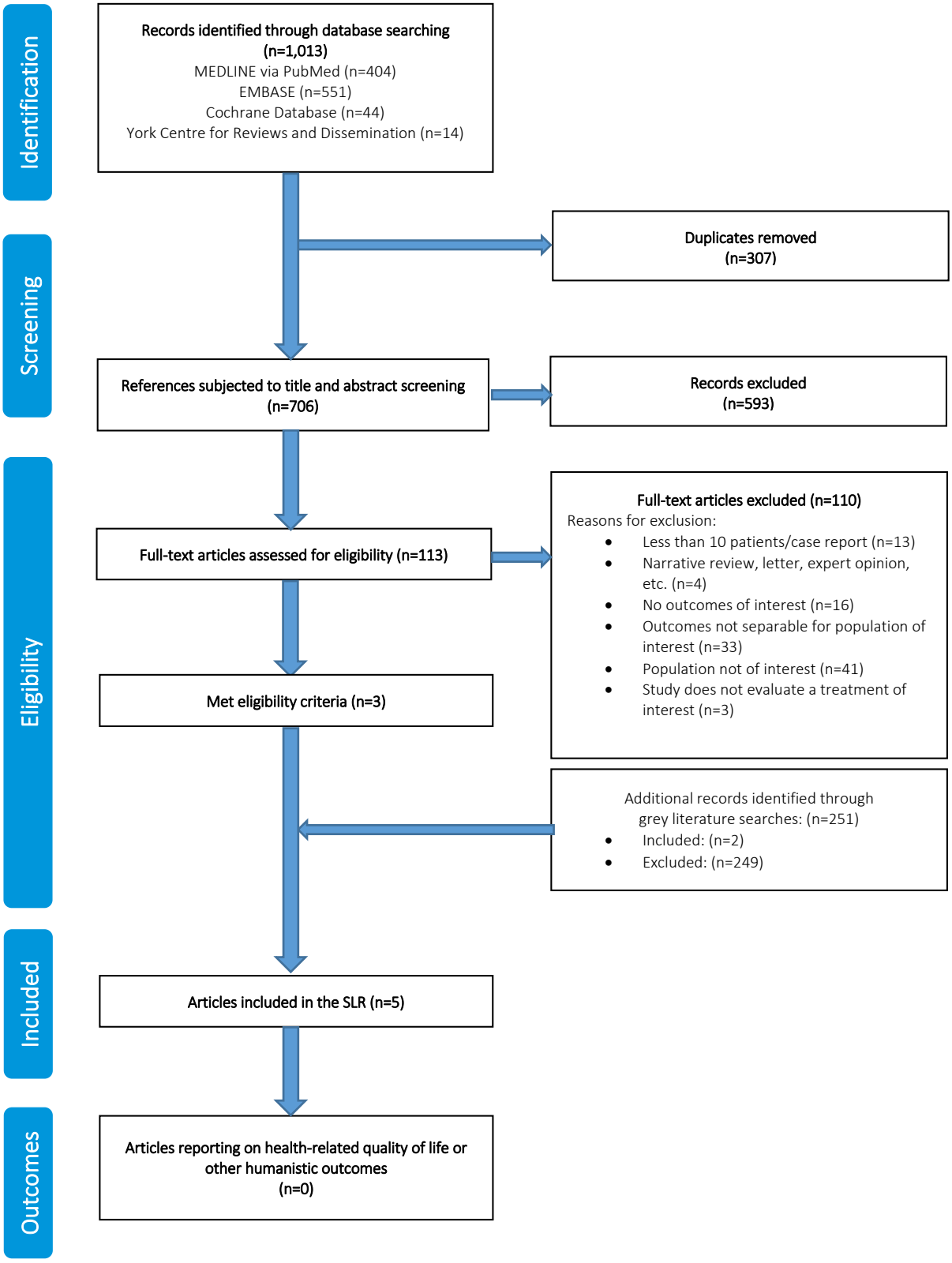
	Inclusion Criteria	Exclusion Criteria
Population	Adult ALL patients with MRD-positivity[i] after treatment[ii]	• Paediatric patients • MRD- ALL patients
Intervention/Comparator	Any interventional therapies	None
Outcomes	Validated measures of HRQoL[iii]	Non-HRQoL outcomes
Study Design	• RCTs of at least 10 patients per arm • Single-arm clinical trials of at least 10 patients • Prospective and retrospective observational studies of at least 10 patients	• Case studies and studies evaluating fewer than 10 patients • Letters, narrative reviews, expert opinions, etc.

Footnotes: [i]Definition of MRD was captured as reported by each study author. It is noted that a number of different technologies are available to measure MRD, with varying specificity and sensitivity. The most common methods include: multicolour flow cytometry to detect abnormal immunophenotypes, with a sensitivity of 10^{-3} to 10^{-4} for 3 to 4 colour flow cytometry and 10^{-4} to 10^{-5} for 6 to 9 color flow cytometry; RT-qPCR assays to detect clonal rearrangements in immunoglobulin (Ig) heavy chain genes, and/or TCR genes, with a sensitivity of 10^{-4} to 10^{-5}; RT-qPCR assays to detect fusion genes (e.g., BCR–ABL), with a sensitivity of 10^{-4} to 10^{-6}. [ii]For inclusion, studies must have involved adult ALL patients who are MRD+ after treatment; data for MRD- populations was extracted from any comparative studies that presented data for both groups. [iii]Observational studies only.
Abbreviations: SLR: systematic literature review; ALL: acute lymphoblastic leukaemia; MRD: minimal residual disease; HRQoL: health-related quality of life; RCT: randomised clinical trial.

Flow diagram for health-related quality of life studies

The following PRISMA diagram, presented in Figure 50, shows the SLR process including the total number of records identified in the searches, and the reasons for study exclusion for health-related quality of life studies. A complete list of excluded publications is provided in Appendix M. No studies providing health-related quality of life evidence were identified in the SLR.

Figure 50. PRISMA diagram for health-related quality of life studies



Abbreviations: PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; SLR: systematic literature review; ALL: acute lymphoblastic leukaemia; MRD: minimal residual disease.

Appendix I Cost and healthcare resource identification, measurement and valuation

A single systematic literature review was performed to identify the available evidence on treatment patterns, humanistic and economic burden, clinical and HRQoL outcomes, and economic evidence relating to the treatment of adults with MRD+ BCP-ALL. The methodology of this SLR is described in detail in Appendix D.

Study Selection

Study selection was performed as described in Appendix D. Table 81 summarises the eligibility criteria used in the study selection process for the costs and healthcare resource use studies.

Table 81. Eligibility criteria used in the study selection process in the costs and healthcare resource use SLR

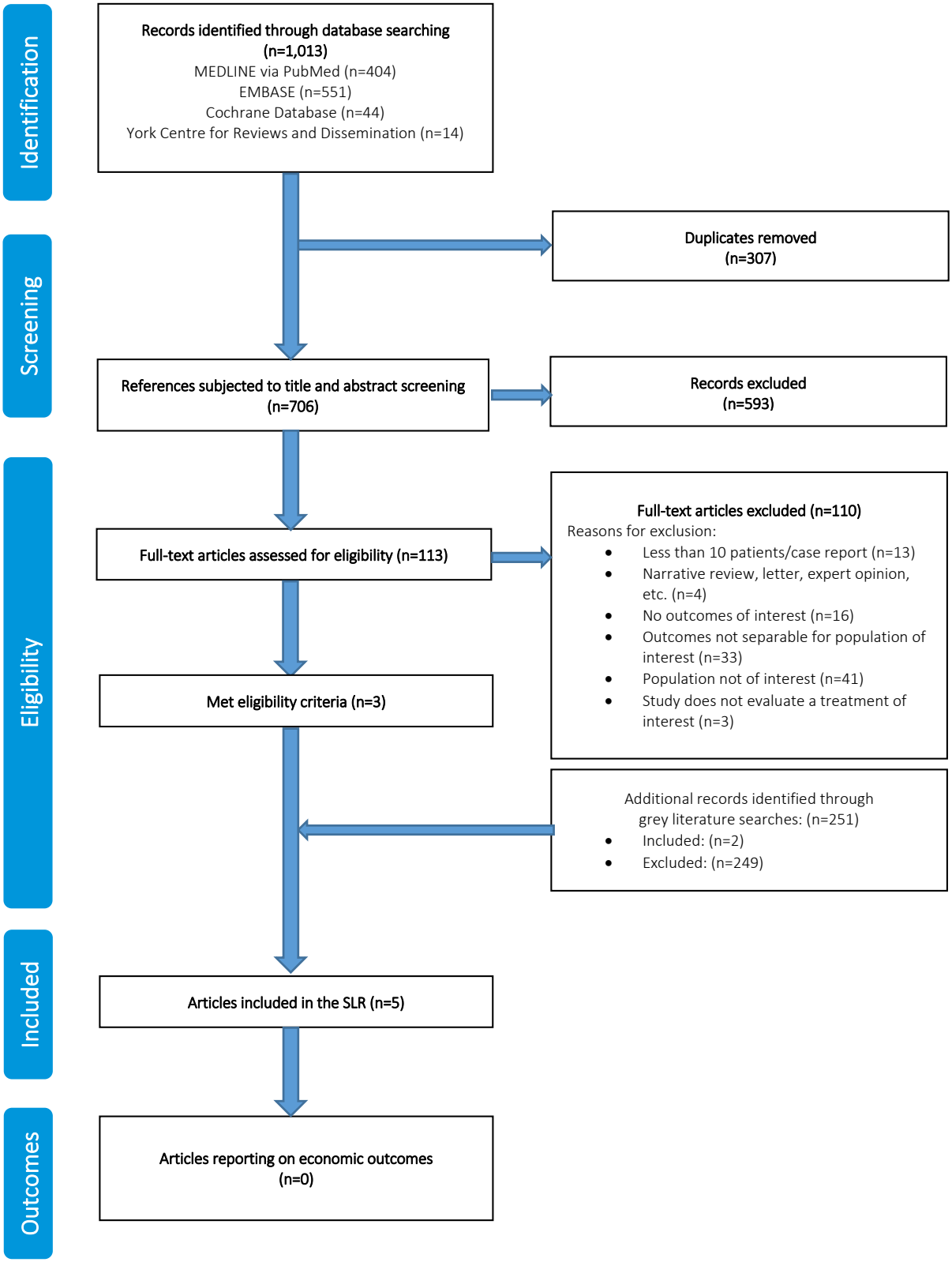
	Inclusion Criteria	Exclusion Criteria
Population	Adult ALL patients with MRD-positivity[i] after treatment[ii]	• Paediatric patients • MRD- ALL patients
Intervention/Comparator	Any interventional therapies	None
Outcomes	Economic burden • Costs (direct and indirect)[iii] • HCRU[iii]	Non-economic outcomes
Study Design	• Economic analyses and HTA reports	Non-economic study designs

Footnotes: [i]Definition of MRD was captured as reported by each study author. It is noted that a number of different technologies are available to measure MRD, with varying specificity and sensitivity. The most common methods include: multicolour flow cytometry to detect abnormal immunophenotypes, with a sensitivity of 10^{-3} to 10^{-4} for 3 to 4 colour flow cytometry and 10^{-4} to 10^{-5} for 6 to 9 color flow cytometry; RT-qPCR assays to detect clonal rearrangements in immunoglobulin (Ig) heavy chain genes, and/or TCR genes, with a sensitivity of 10^{-4} to 10^{-5}; RT-qPCR assays to detect fusion genes (e.g., BCR–ABL), with a sensitivity of 10^{-4} to 10^{-6}. [ii]For inclusion, studies must have involved adult ALL patients who are MRD+ after treatment; data for MRD- populations was extracted from any comparative studies that presented data for both groups. [iii]Observational studies only.
Abbreviations: SLR: systematic literature review; ALL: acute lymphoblastic leukaemia; MRD: minimal residual disease; HCRU: healthcare resource utilisation; HTA: health technology assessment.

Flow diagram for cost and healthcare resource use studies

The following PRISMA diagram, presented in Figure 51, shows the SLR process including the total number of records identified in the searches, and the reasons for study exclusion for cost and healthcare resource use studies. A complete list of excluded publications is provided in Appendix M. No studies providing cost and healthcare resource use evidence were identified in the SLR.

Figure 51. PRISMA diagram for cost and healthcare resource use



Abbreviations: PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; SLR: systematic literature review; ALL: acute lymphoblastic leukaemia; MRD: minimal residual disease.

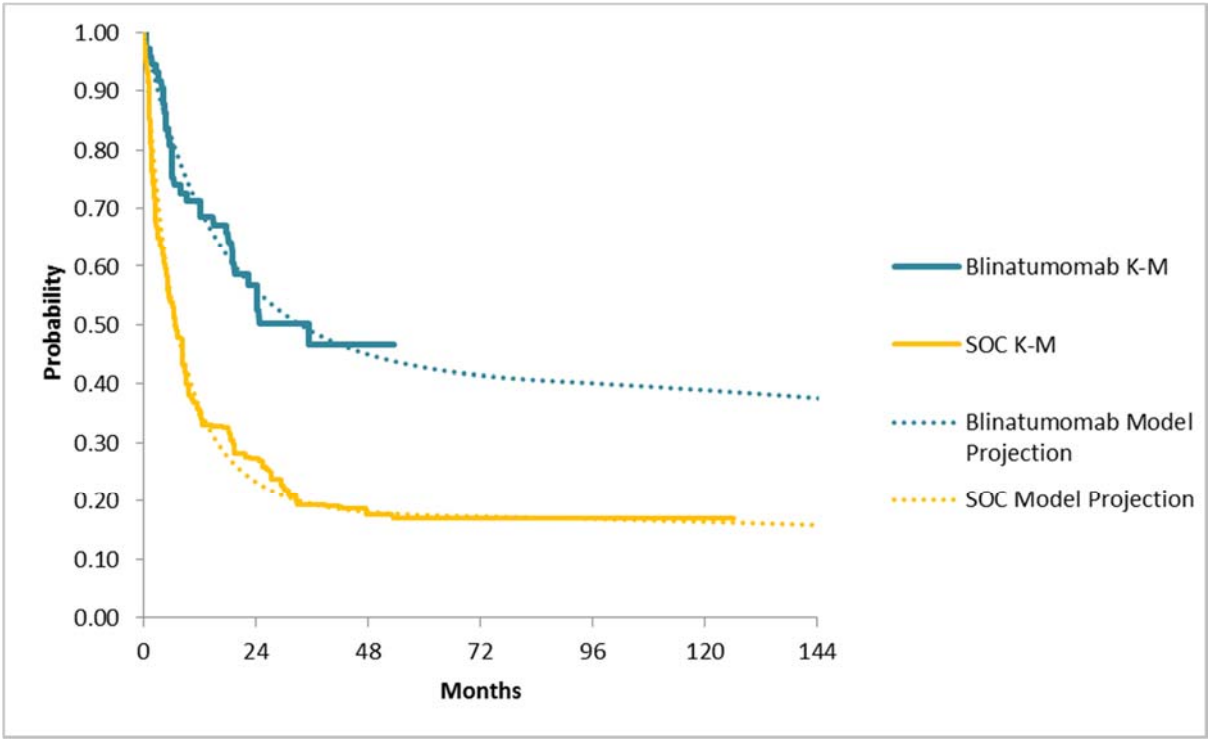
Appendix J Clinical outcomes and disaggregated results from the model

a. Clinical outcomes from the model

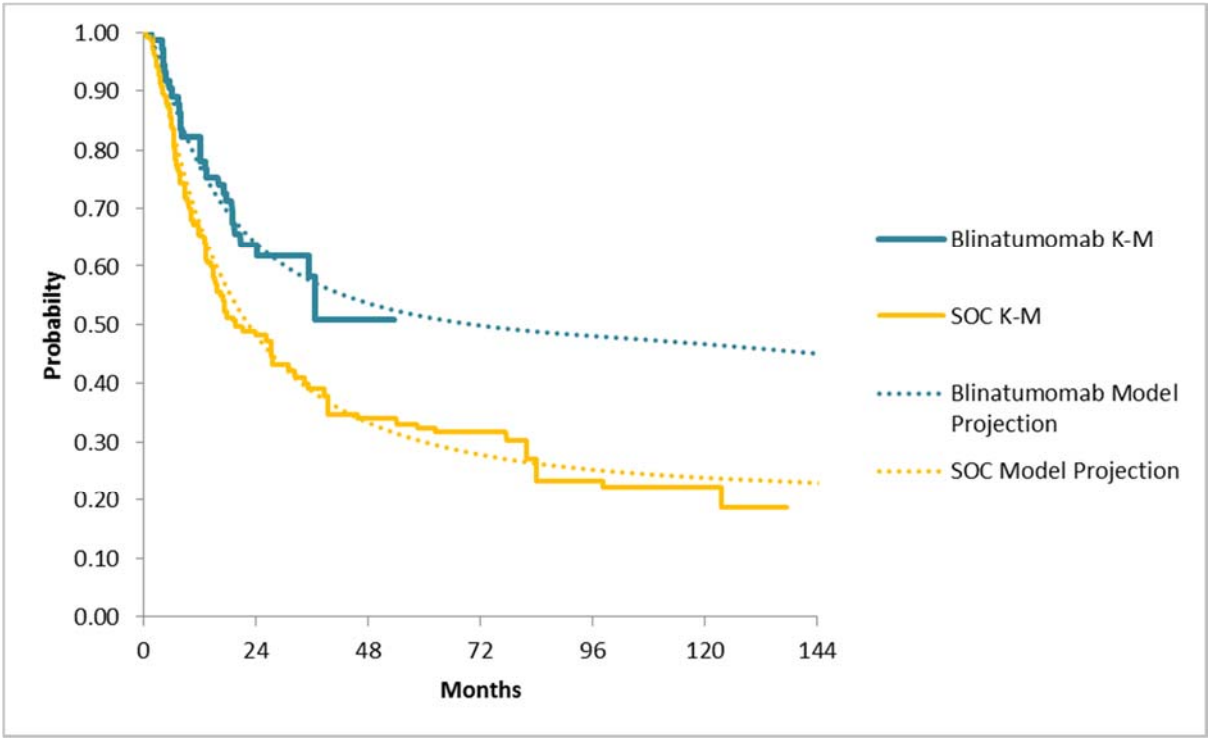
The main clinical outcomes generated by the model are RFS and OS. Estimates of RFS and OS from the model are compared with K-M estimates of RFS and OS from BLAST and the historical control study (based on the patients in the propensity matched analysis and using IPTW-ATT weights for the historical control) in Figure 52 and Table 82. At 53.5 months (the last observed failure or censoring time in BLAST), the model projections very closely approximate the K-M survival probabilities for blinatumomab and SoC.

Figure 52. Relapse-free survival and overall survival in the model

Relapse-free survival



Overall survival



Footnotes: Model time horizon is 50 years in base-case.
Abbreviations: K-M: Kaplan-Meier; SoC: standard of care.

Table 82. Comparison of probabilities of survival in the model and in BLAST at selected landmarks

Relapse-free survival

	Blinatumomab		SoC chemotherapy	
Month	BLAST	Model	Historical control	Model
6	76.0%	81.4%	52.8%	52.9%
12	69.2%	69.1%	34.4%	35.0%
24	55.7%	55.4%	27.3%	23.0%
53.5	46.6%	43.7%	17.0%	17.7%

Overall survival

	Blinatumomab		SoC chemotherapy	
Month	BLAST	Model	Historical control	Model
6	89.0%	88.4%	83.4%	83.4%
12	79.5%	76.9%	65.3%	66.8%
24	62.8%	63.8%	48.2%	48.0%
53.5	50.9%	52.4%	34.0%	31.6%

Abbreviations: SoC: standard of care.

A breakdown of life-years and QALYs by state is shown in Table 83. Blinatumomab yields 3.58 more discounted life-years and 2.95 more discounted QALYs than SoC.

Table 83. Base-case effectiveness results

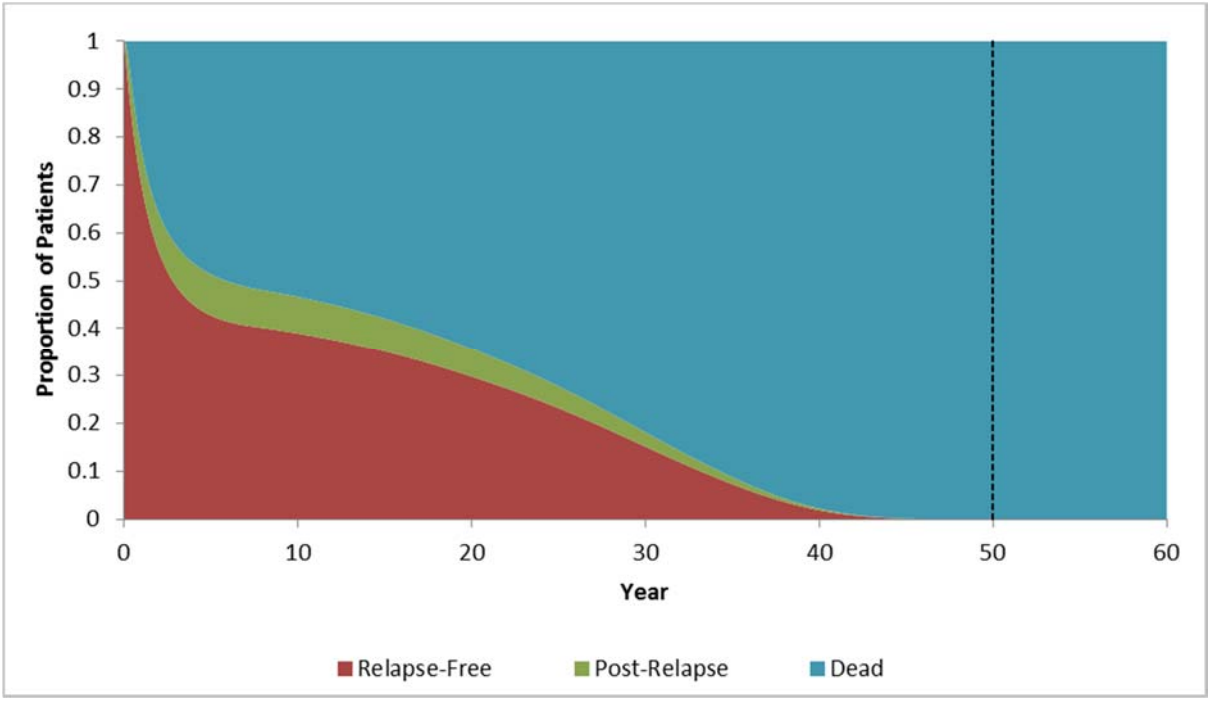
Effectiveness	Blinatumomab	SoC	Incremental	Absolute Incremental	Absolute Incremental %
Undiscounted					
Life-years					
Relapse-free	11.50	5.06	6.43	6.43	112.1
Post-relapse	2.10	2.79	-0.70	0.70	12.1
Total	13.59	7.86	5.74	5.74	100.0
QALYs					
Relapse-free	8.99	3.92	5.08	5.08	110.0
Post-relapse	1.55	2.01	-0.46	0.46	10.0
Total	10.54	5.93	4.61	4.61	100.0
Discounted					
Life-years					
Relapse-free	7.73	3.48	4.25	4.25	118.8
Post-relapse	1.36	2.03	-0.67	0.67	18.8
Total	9.09	5.51	3.58	3.58	100.0
QALYs					
Relapse-free	6.11	2.71	3.40	3.40	115.1
Post-relapse	0.99	1.44	-0.45	0.45	15.1
Total	7.10	4.14	2.95	2.95	100.0

Abbreviations: SoC: standard of care; QALY: quality-adjusted life year.

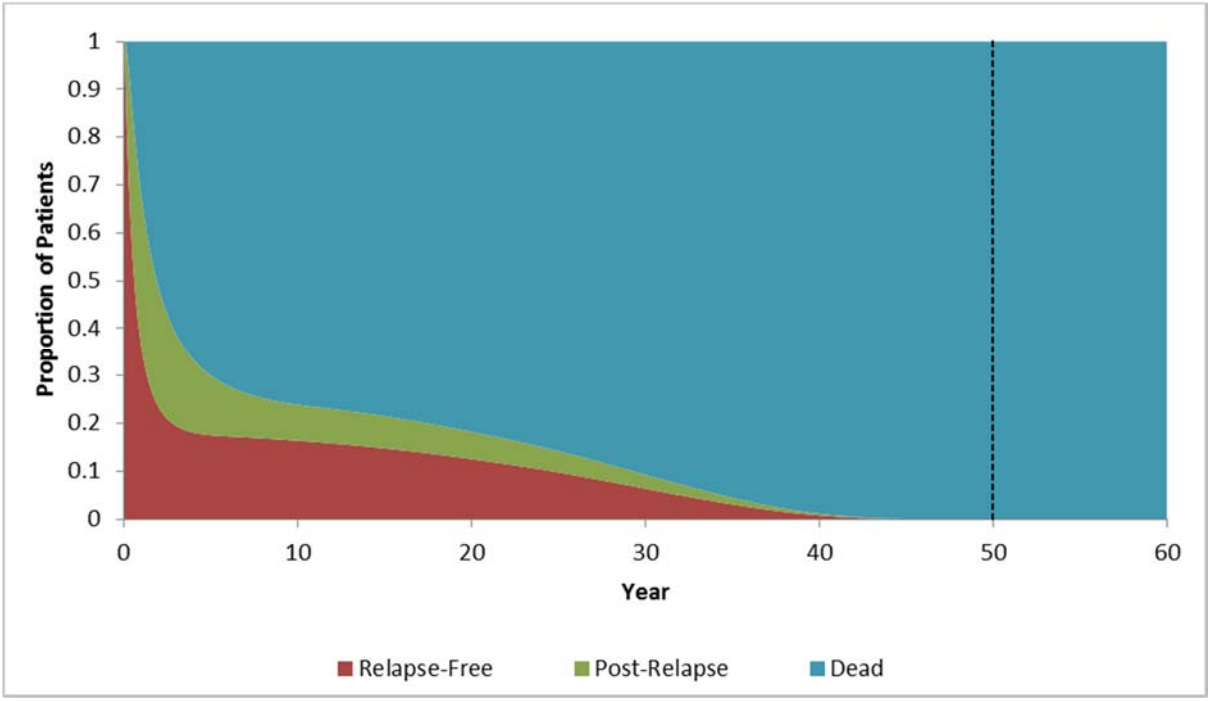
Survival traces showing the proportion of patients in each state for blinatumomab and SoC are shown in Figure 53.

Figure 53. Survival trace for blinatumomab (top) and SoC (bottom)

Blinatumomab



SoC



Abbreviations: SoC: standard of care.

Table 84. Landmark RFS and OS for blinatumomab and SoC maintenance therapy from the model compared with results from IPTW-ATT propensity matched comparison of BLAST and the historical control study

Month	Blinatumomab	SoC

	BLAST	Model	BLAST	Model
RFS				
6	76.0%	81.4%	52.8%	52.9%
12	69.2%	69.1%	34.4%	35.0%
24	55.7%	55.4%	27.3%	23.0%
53.5	46.6%	43.7%	17.0%	17.7%
OS				
6	89.0%	88.4%	83.4%	83.4%
12	79.5%	76.9%	65.3%	66.8%
24	62.8%	63.8%	48.2%	48.0%
53.5	50.9%	52.4%	34.0%	31.6%

Abbreviations: RFS: relapse-free survival; OS: overall survival; SoC: standard of care; IPTW: inverse probability of treatment weighting; ATT: average treatment effect on the treated

b. Disaggregated results of the base-case incremental cost-effectiveness analysis

Base-case results for cost outcomes are shown in Table 85. Medication costs were estimated to be £83,129 higher with blinatumomab versus SoC. Total treatment costs, including medication, hospitalisation, outpatient visits, and infusion pump costs were estimated to be £88,788 higher with blinatumomab. Costs of pre-relapse HSCT were £36,552 higher with blinatumomab than SoC. Other pre-relapse inpatient costs were estimated to be £11,275 lower with blinatumomab versus SoC, whereas other pre-relapse outpatient costs were £3,830 higher with blinatumomab versus SoC. Projected other inpatient costs were higher with SoC because the difference in estimated inpatient costs for patients with versus without MRD response is sufficiently large to offset the longer time in RFS among blinatumomab patients. Conversely, for other outpatient costs, the difference in costs for patients with versus without MRD response *was not* sufficiently large to offset the longer time in RFS among blinatumomab patients.

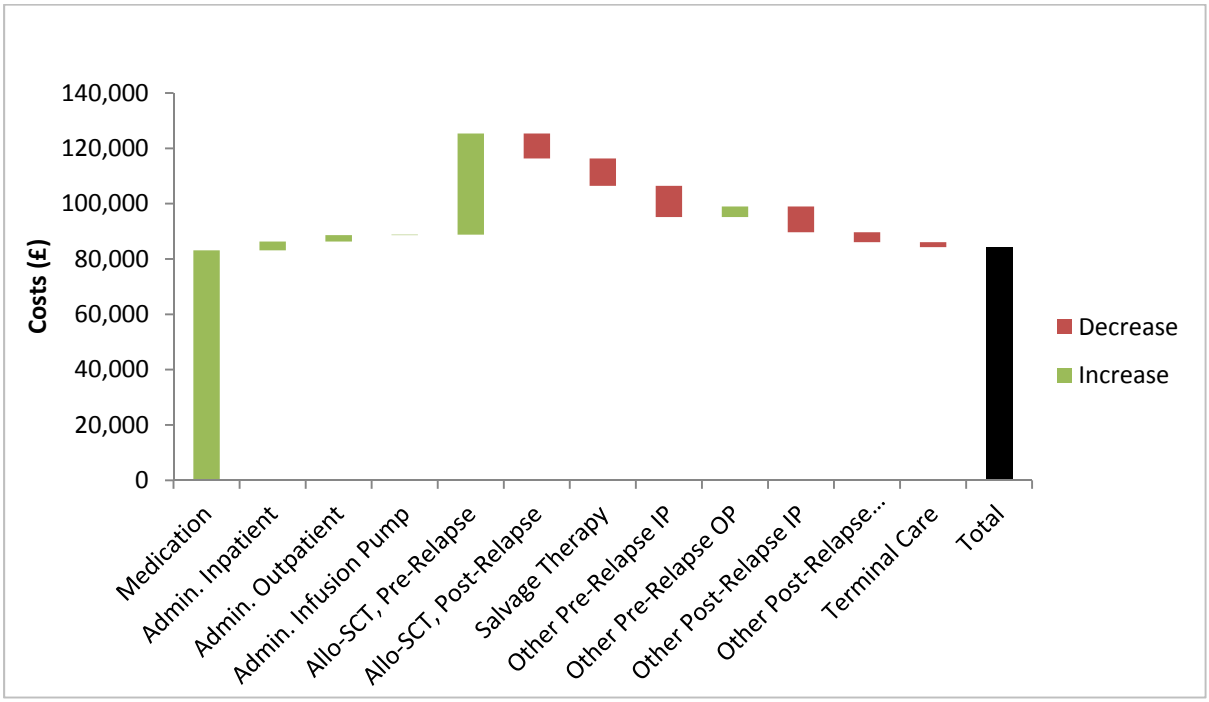
Because fewer patients receiving blinatumomab are projected to relapse, salvage therapy costs were £9,905 higher with SoC than with blinatumomab. Post-relapse allogeneic HSCT costs were £9,024 higher with SoC than with blinatumomab. Because post-relapse LYs were projected to be greater with SoC than with blinatumomab, other post-relapse inpatient costs were £9,306 higher with SoC than with blinatumomab. Similarly, other post-relapse outpatient costs were £3,576 higher with SoC than with blinatumomab. Total post-relapse costs were £31,811 higher with SoC than blinatumomab. Because fewer blinatumomab patients are projected to die within 5 years, terminal care costs were £1,825 higher for SoC than for blinatumomab. Total incremental costs were £84,259 higher with blinatumomab versus SoC. A waterfall diagram of incremental costs with blinatumomab versus SoC is shown in Figure 54.

Table 85. Summary of predicted resource use by category of cost

	Blinatumomab (£)	SoC (£)	Incremental (£)	Absolute incremental (£)	Absolute incremental (%)
Blinatumomab and SoC maintenance treatment					
Medication	84,174	1,046	83,129	83,129	98.7
Administration					
Hospitalisation	3,185	N/A	3,185	3,185	3.8
Outpatient visits	3,645	1,354	2,291	2,291	2.7
Infusion pump	182	N/A	182	182	0.2
Total medication administration	91,187	2,399	88,788	88,788	105.4
HSCT	75,266	38,714	36,552	36,552	43.4
Other inpatient, pre-relapse	7,046	18,321	-11,275	11,275	13.4
Other outpatient, pre-relapse	11,278	7,449	3,830	3,830	4.5
Post-relapse					
Salvage therapy	6,450	16,355	-9,905	9,905	11.8
HSCT	7,678	16,701	-9,024	9,024	10.7
Other inpatient	5,276	14,582	-9,306	9,306	11.0
Other outpatient	2,036	5,612	-3,576	3,576	4.2
Total post-relapse	21,439	53,250	-31,811	31,811	37.8
Terminal Care	4,167	5,992	-1,825	1,825	2.2
Total	210,383	126,124	84,259	84,259	100.0

Abbreviations: SoC: standard of care; HSCT: allogeneic haematopoietic stem cell transplant.

Figure 54. Waterfall diagram of incremental costs with blinatumomab versus SoC



Abbreviations: SoC: standard of care; allo-SCT: allogeneic haematopoietic stem cell transplantation; IP: inpatient; OP: outpatient.

Resource utilisation was not explicitly tallied in the model.

Appendix K Checklist of confidential information

The checklist of confidential information is included as a separate document alongside this submission.

Appendix L Propensity Score Analysis Methodology and ATE Results

Analysis sets

The Primary Analysis Set for the PS analysis included patients who adhered to the following criteria:

BLAST (MT103-203) criteria:

- Received any infusion of the investigational drug, blinatumomab
- Philadelphia negative B–precursor ALL in complete haematological remission defined as less than 5% blasts in bone marrow after at least three intensive chemotherapy blocks
- MRD+ at a level of ≥1 x 10^{-3} (PCR only in BLAST) but otherwise in complete haematological remission
- At least 18 years old at the MRD baseline date
- In their first haematological remission (CR1 only)

Historical comparator study (20120148) criteria:

- Ph-negative B-precursor ALL in complete haematological remission
- MRD+ at a level of ≥1 x 10^{-3} regardless of detection method
- At least 18 years old at the MRD baseline date
- Time to relapse greater than 14 days from the date of MRD detection (see explanation below)
- In their first haematological remission (CR1 only)

To correctly compare RFS and OS, a baseline date must be well aligned between the two populations. Initially two different baseline dates were used in the trials: the date of first blinatumomab treatment for BLAST patients, and the date of MRD detection for the historical comparator patients. However, an immortal time bias is introduced by using these dates because historical comparator patients with a rapid relapse following MRD detection would not have counterparts in the BLAST study. To better align the populations and reduce bias due to the definition of MRD baseline date, historical comparator patients were excluded if their time to relapse was less than 14 days, which is the median time between MRD detection and first blinatumomab dose for BLAST patients. In addition, the baseline date for historical comparator patients was set at MRD detection date plus 14 days.

Defining MRD baseline detection date

To correctly compare RFS and OS, a baseline date must be well aligned between the two populations. Aligning both populations using their MRD detection date would lead to an immortal time bias for MT103-203 patients, due to the fact that patients relapsing or dying after MRD detection but before beginning treatment with blinatumomab would not have been included in the MT103-203 study. Initially two different baseline dates were used: the date of first blinatumomab treatment for MT103-203 patients, and the date of MRD detection for 20120148 patients. However, an immortal time bias is introduced by in using these dates because 20120148

patients with quick relapse following MRD detection would not have counterparts in the MT103-203 study. To better align the populations and reduce bias due to the definition of MRD baseline date, study 20120148 historical patients were excluded if their time to relapse was less than 14 days, which is the median time between MRD detection and first blinatumomab dose for MT103-203 patients. In addition, the baseline date for historical control patients from the 20120148 study was set at MRD detection date plus 14 days.

The baseline date (i.e. start time) for relapse-free and overall survival was therefore defined as 14 days after the MRD baseline date for 20120148 patients and the date of the first blinatumomab treatment for MT103-203 patients. Because study 20120148 captures an extended disease history, some patients might have had multiple MRD detection dates following multiple complete remissions from multiple lines of chemotherapy; at each of these dates the patient would have been eligible for study MT103-203 (provided they had at least three total blocks of chemotherapy). For these patients, the MRD baseline date was defined as 14 days after the date of first MRD detection following the first complete remission (if data were available).

Missing Data

There was a limited amount of missing data in the covariate set for the propensity score analysis. Only two variables contained missing values. Two patients were missing the age at MRD baseline, which was used as inclusion criteria. These individuals had values for their age at diagnosis which were much greater than 18, therefore it was safely inferred that they met the age requirement.

There were approximately 20 individuals in study MT103-203 who did not have a white blood cell count value at diagnosis. For these patients, multiple imputation was applied via PROC MI in SAS, which created a single set of imputed values for the categorical white blood cell count (<=30,000/µl, >30,000/µl) using a logistic regression model based on all remaining baseline covariates.

Data Synthesis and Statistical Analysis

The databases from the two studies were merged programmatically and used for analysis. Once combined the following steps were carried out to complete the statistical analysis:

- Select candidate variables for the propensity score model. Candidate variables are those that are common to both the databases and are thought to be important for characterizing the blinatumomab treated population. Candidate variables were selected based on their prognostic potential determined through study team discussions.

- Run the variable selection algorithm in order to choose the variables and interaction terms considered relevant for discriminating between those who were and were not treated with blinatumomab. The final model is used for generating each subject's propensity score.

- Evaluate the propensity score overlap between treatment groups via a box plot and evaluate the balance between treatment groups before and after propensity score (PS) adjustments.

- If balance is adequately achieved, conduct the endpoint analyses (RFS and OS) using the appropriate inverse probability of treatment (IPT) weights.

Details of these steps are outlined in the following sections.

Propensity Score Model Development

Candidate variable main effects and all two-way interaction terms were entered into a logistic regression model with blinatumomab treatment as the binary response. A stepwise variable selection algorithm was run whereby p<0.30 was used as the threshold for entering and keeping covariates in the model.

- Candidate propensity score model covariates included the following:
- Age at primary diagnosis (years)
- Sex (male, female)
- Country (Germany, others)
- Presence and type of any cytogenetic and molecular aberrations (t(4;11)MLL-AF4 (Yes, No/Unknown)
- Time from primary diagnosis to MRD baseline date (months)
- Baseline MRD level (-3="<1x10—3", -2= ">=1x10^{-3}and <1x10^{-2}" , -1= ">=1x10^{-2} and <1x10^{-1}" , 0= ">=1x10^{-1}")
- White blood cells (WBC) at diagnosis (<=30,000/µl, >30,000/µl)
- Type of prior chemotherapy (GMALL, other)
- Baseline MRD was recoded into an ordinal variable as defined above, and was treated as a continuous covariate in the model.
- A propensity score model was fit for each analysis set separately.

Balance Diagnostics

Upon deriving propensity scores for each patient, balance between the two treatment groups with respect to their PS was assessed via box plots. The overall balance was to be considered sufficient if at least 25% of the historical data overlapped with the inner 95th percentile of the blinatumomab data, as pre-specified in the SSAP.

With respect to individual covariates considered for the propensity score model, two methods were employed to ascertain the balance between the data sources before and after propensity score adjustments. The first method involved univariate regression models with the baseline factor as the dependent variable and the treatment group as the independent variable. For categorical factors, a logistic regression model with robust variance estimation was used. For continuous variables, a general linear model with robust variance estimation was used. The p-value associated with the treatment group effect from each model was used to compare the before- and after-effects of the PS adjustment.

The second method involved calculation of standardized differences. Standardized differences can be used to ascertain the balance in a way that is not dependent on the sample size. For a continuous variable, the standardized differences were calculated as:

$$d = \frac{(\overline{X_1} - \overline{X_2})}{\sqrt{\frac{s_1^2 + s_2^2}{2}}}$$

For a binary variable, the standardized differences were calculated as:

$$d = \frac{(\widehat{P_1} - \widehat{P_2})}{\sqrt{\frac{[\widehat{P_1}(1-\widehat{P_1}) + \widehat{P_2}(1-\widehat{P_2})]}{2}}}$$

Criteria for deciding whether the balance was adequate included: univariate p-values greater after adjustment and not considered significant and a standardized difference less than at least 0.20 (best balance achieved when less than 0.10).

If important covariates or baseline factors were not adequately balanced after conducting the evaluations described above and the covariate was considered prognostic with respect to the endpoint, then those factors may have been added as additional covariates to the endpoint analysis model for sensitivity analyses.

Propensity Score Adjustment Method

The inverse probability of treatment weighting (IPTW) approach for propensity score adjustment was used for this analysis. With time-to-event endpoints, stratification and covariate adjustment have been shown to produce biased estimates of marginal and conditional hazard ratios, and the

limited sample size prohibits matching. Different weights can be applied depending on the objective of the analysis and the methodology is discussed below.

Average treatment effect of the treated weights were also considered (ATT)

$$w_i = Z_i + \frac{e_i(1 - Z_i)}{(1 - e_i)}$$

Where the subscript i denotes the ith subject, w represents the weight, Z is assigned a value of 1 for treated (blinatumomab) subjects and 0 for untreated subjects, and e represents the propensity score.

Average treatment effect (ATE)

$$w_i = \frac{Z_i}{e_i} + \frac{(1 - Z_i)}{(1 - e_i)}$$

Where the subscript i denotes the ith subject, w represents the weight, Z is assigned a value of 1 for treated (blinatumomab) subjects and 0 for untreated subjects, and e represents the propensity score.

Results (ATE weighting)

Balance in baseline covariates using ATE weighting

Table 86 summarises the degree of imbalance between treatment groups with respect to the 9 covariates prior to and after adjustment. Prior to adjustment, the covariates with the largest difference between treatment groups were age at diagnosis (mean 36.3 and 44.8 for historical control and blinatumomab populations, respectively) and time from diagnosis to baseline (mean 6.6 and 12.8 for historical control and blinatumomab populations, respectively). Additional covariates, with a standardised difference greater than 0.2, were country, WBC at diagnosis (continuous), and prior chemotherapy. Additional covariates, with a standardised difference greater than 0.1, were MRD at baseline and WBC at diagnosis (categorical). After adjustment, none of the p-values were significant and 8 of the 9 covariates had standardised differences less than 0.2, and 6 out of 9 had standard differences less than 0.1. In both cases, the standardised difference values were borderline. Therefore, inclusion of these covariates in the outcome model was not warranted.

Table 86. Covariate Balance Before and After Propensity Score Adjustments using ATE weighting

Characteristic	Unweighted				Stabilised IPTW			
Mean (SD)/n (%)	Control (N=182)	Blinatumomab (N=73)	Standard Difference	P-value	Control (N=174.3)	Blinatumomab (N=78.5)	Standard Difference	P-value
Age at primary diagnosis (years)	36.3 (13.6)	44.8 (16.6)	-0.56	<0.001	37.8 (13.8)	36.5 (16.4)	0.09	0.573
Gender (Female)	80 (44.0)	32 (43.8)	0.00	0.986	76.6 (43.9)	27.0 (34.4)	0.20	0.226
Country (Not Germany)	112 (61.5)	35 (47.9)	0.28	0.048	143.6 (58.8)	151.6 (55.3)	0.07	0.674
MRD at Baseline (recoded)	-1.5 (0.6)	-1.7 (0.7)	0.16	0.249	-1.6 (0.60)	-1.5 (0.8)	-0.08	0.688
Time from diagnosis to baseline (months)	6.6 (6.1)	12.8 (14.3)	-0.56	<.001	7.3 (7.2)	8.1 (9.7)	-0.09	0.463
WBC at diagnosis (>30,000/mm^3)	51 (28.0)	15 (20.5)	0.17	0.220	45.2 (26.0)	19.1 (24.3)	0.04	0.822
WBC at diagnosis (continuous, log10)	4.15 (0.62)	3.98 (0.60)	0.26	0.072	4.13 (0.60)	4.07 (0.60)	0.10	0.542
T411mll4 mutation (Yes)	15 (8.2)	5 (6.8)	0.05	0.709	14.1 (8.1)	5.6 (7.2)	0.03	0.820
Prior chemotherapy (GMALL)	76 (41.8)	42 (57.5)	-0.32	0.023	78.0 (44.7)	39.2 (50.0)	-0.10	0.533

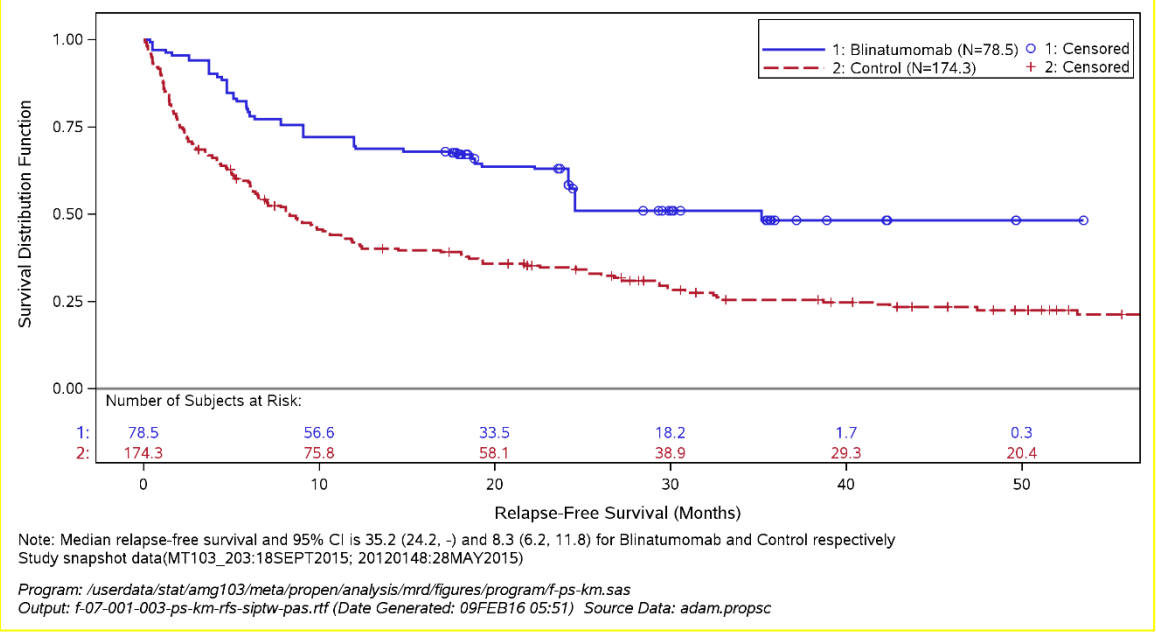
Abbreviations: IPTW: inverse probability of treatment weighting; SD: standard deviation; MRD: minimal residual disease; WBC: white blood cell; GMALL: German Multicentre ALL Working Group.

RFS

- **Compared to SoC, blinatumomab provides a statistically significant reduction in the risk of relapse or death of 50%.**

In the analysis without adjusting for HSCT, 18-month RFS was 0.39 (95% CI: 0.33, 0.48) for control and 0.67 (95% CI: 0.58, 0.78) for blinatumomab, representing a 1.7-fold increase in 18-month RFS, as presented in Figure 55. K-M based median RFS (95% CI), unadjusted for HSCT, was estimated at 8.3 months (6.2, 11.8) for control and 35.2 months (24.2, n.e.) for blinatumomab representing a 26.9-month improvement in median RFS. The K-M curves demonstrate a clear separation in relapse-free survival over time between the two treatment groups, and these results demonstrate a statistically significant association between blinatumomab treatment and improvements in RFS compared to SoC. When adjusting for HSCT, blinatumomab resulted in statistically significantly longer RFS, with a 50% reduction in the risk of relapse or death compared to historical controls (HR: 0.50; 95% CI: 0.32, 0.78).[65]

Figure 55. RFS in BLAST versus SoC chemotherapy using ATE weighting, unadjusted



Note: Median relapse-free survival and 95% CI is 35.2 (24.2, -) and 8.3 (6.2, 11.8) for Blinatumomab and Control respectively
Study snapshot data(MT103_203:18SEPT2015; 20120148:28MAY2015)

Program: /userdata/stat/amg103/meta/propen/analysis/mrd/figures/program/f-ps-km.sas
Output: f-07-001-003-ps-km-rfs-siptw-pas.rtf (Date Generated: 09FEB16 05:51) Source Data: adam.propsc

Abbreviations: RFS: relapse-free survival; SoC: Standard of Care; ATE: average treatment effect, population level.

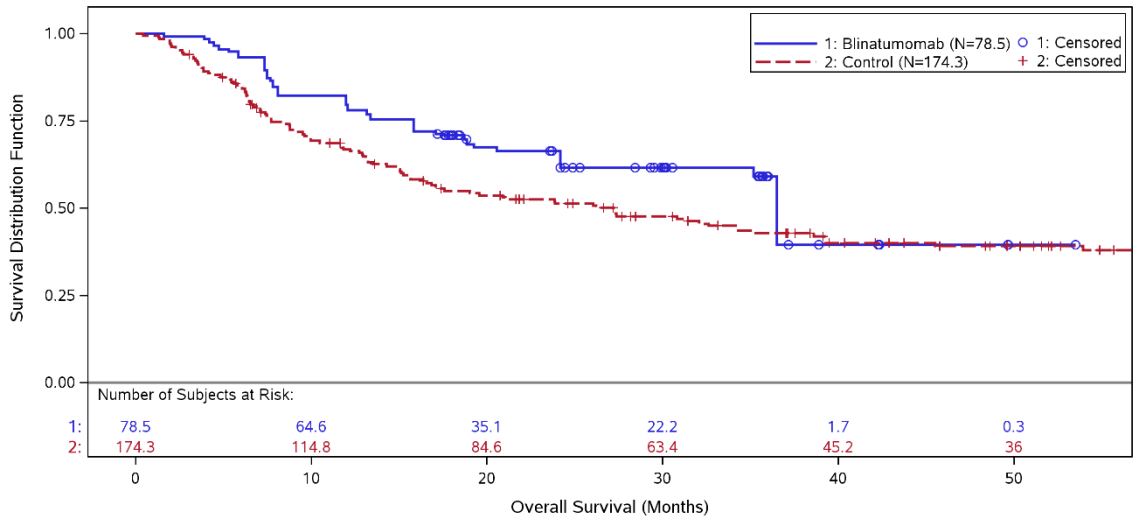
OS

- **Compared to SoC, blinatumomab provides a numerical improvement in OS.**

In the analysis without adjusting for HSCT, 18-month OS at was 0.55 (95% CI: 0.48, 0.63) for control and 0.71 (95% CI: 0.62, 0.81) for blinatumomab, representing a 1.3-fold increase in 18-month OS, as presented in Figure 56. K-M based median OS (95% CI), unadjusted for HSCT, was estimated at 27.2 months (16.4, 38.6) for control and 36.5 months (24.2, n.e.) for blinatumomab, representing a 9.3-month improvement in median OS. These results suggest a directional improvement in OS due to blinatumomab compared to SoC. The lack of statistical significance in this analysis could be attributed to the fewer number of deaths in the data set,

compared to relapses, which results in less power and wider confidence intervals. When adjusting for HSCT, a directional but not numerically significant improvement associated with blinatumomab compared to historical controls was observed.[65]

Figure 56. OS in BLAST versus SoC chemotherapy using ATE weighting, unadjusted



Note: Median overall survival and 95% CI is 36.5 (24.2, -) and 27.2 (16.4, 38.6) for Blinatumomab and Control respectively
Study snapshot data(MT103_203:18SEPT2015; 20120148:28MAY2015)

Program: /userdata/stat/amg103/meta/propen/analysis/mrd/figures/program/f-ps-km.sas
Output: f-08-001-003-ps-km-os-siptw-pas.rtf (Date Generated: 09FEB16 05:51) Source Data: adam.propsc

Abbreviations: OS: Overall survival; SoC: Standard of Care; ATE: average treatment effect, population level.

Appendix M Articles excluded from the SLR

A complete list of studies excluded from the SLR (including clinical, cost-effectiveness, health-related quality of life, and cost and healthcare resource use studies) after the full text review stage, with reasons for exclusion, is included in Table 87.

Table 87. Complete list of articles excluded after full text review from the SLR

Reference ID	Authors	Title	Reason for Exclusion
6	Pemmaraju N., Kantarjian H., Jorgensen J. L., Jabbour E., Jain N., Thomas D., O'Brien S., Wang X., Huang X., Wang S. A., Konopleva M., Konoplev S., Kadia T., Garris R., Pierce S., Garcia-Manero G., Cortes J., Ravandi F.	Significance of recurrence of minimal residual disease detected by multi-parameter flow cytometry in patients with acute lymphoblastic leukemia in morphological remission	Population not of interest
9	Kebriaei P., Bassett R., Lyons G., Valdez B., Ledesma C., Rondon G., Oran B., Ciurea S., Alousi A., Popat U., Patel K., Ahmed S., Olson A., Bashir Q., Shah N., Jones R., Marin D., Rezvani K., Nieto Y., Khouri I., Qazilbash M., Hosing C., Shpall E., Champlin R. E., Andersson B. S.	Clofarabine Plus Busulfan is an Effective Conditioning Regimen for Allogeneic Hematopoietic Stem Cell Transplantation in Patients with Acute Lymphoblastic Leukemia: Long-Term Study Results	Population not of interest
13	Motlló C., Ribera J. M., Morgades M., Granada I., Montesinos P., Brunet S., Bergua J., Tormo M., García-Boyero R., Sarrà J., del Potro E., Grande C., Barba P., Bernal T., Amigo M. L., Grau J., Cervera J., Feliu E.	Frequency and prognostic significance of t(v;11q23)/KMT2A rearrangements in adult patients with acute lymphoblastic leukemia treated with risk-adapted protocols	No outcomes of interest
14	Brammer J. E., Saliba R. M., Jorgensen J. L., Ledesma C., Gaballa S., Poon M., Maziarz R. T., Champlin R. E., Hosing C., Kebriaei P.	Multi-center analysis of the effect of T-cell acute lymphoblastic leukemia subtype and minimal residual disease on allogeneic stem cell transplantation outcomes	Population not of interest
15	Giebel S., Labopin M., Socié G., Beelen D., Browne P., Volin L., Kyrcz-Kr Zemien S., Yakoub-Agha I., Aljurf M., Wu D., Michallet M., Arnold R., Mohty M., Nagler A.	Improving results of allogeneic hematopoietic cell transplantation for adults with acute lymphoblastic leukemia in first complete remission: An analysis from the Acute Leukemia Working party of the European Society for Blood and Marrow Transplantation	Population not of interest
17	Jabbour E., Short N. J., Jorgensen J. L., Yilmaz M., Ravandi F., Wang S. A., Thomas D. A., Khoury J., Champlin R. E., Khouri I., Kebriaei P., O'Brien S. M., Garcia-Manero G., Cortes J. E., Sasaki K., Dinardo C. D., Kadia T. M., Jain N., Konopleva M., Garris R., Kantarjian H. M.	Differential impact of minimal residual disease negativity according to the salvage status in patients with relapsed/refractory B-cell acute lymphoblastic leukemia	Population not of interest
24	Aristei C., Carotti A., Palazzari E., Amico L., Ruggeri L., Perrucci E., Falcinelli L., Lancellotta V., Palumbo I., Falzetti F., Aversa F., Merluzzi M., Velardi A., Martelli M. F.	The Total Body Irradiation Schedule Affects Acute Leukemia Relapse After Matched T Cell–Depleted Hematopoietic Stem Cell Transplantation	Outcomes not separable for population of interest

Reference ID	Authors	Title	Reason for Exclusion
25	Lussana F., Intermesoli T., Gianni F., Boschini C., Masciulli A., Spinelli O., Oldani E., Tosi M., Grassi A., Parolini M., Audisio E., Cattaneo C., Raimondi R., Angelucci E., Cavattoni I. M., Scattolin A. M., Cortelezzi A., Mannelli F., Ciceri F., Mattei D., Borlenghi E., Terruzzi E., Romani C., Bassan R., Rambaldi A.	Achieving Molecular Remission before Allogeneic Stem Cell Transplantation in Adult Patients with Philadelphia Chromosome–Positive Acute Lymphoblastic Leukemia: Impact on Relapse and Long-Term Outcome	Outcomes not separable for population of interest
31	Zhang M., Fu H., Lai X., Tan Y., Zheng W., Shi J., Zhao Y., Lin M., He J., Cai Z., Luo Y., Huang H.	Minimal residual disease at first achievement of complete remission predicts outcome in adult patients with Philadelphia chromosome-negative acute lymphoblastic leukemia	Population not of interest
35	Yan C. H., Wang Y., Wang J. Z., Chen Y. H., Chen Y., Wang F. R., Sun Y. Q., Mo X. D., Han W., Chen H., Zhang X. H., Xu L. P., Liu K. Y., Huang X. J.	Minimal residual disease- and graft-vs.-host disease-guided multiple consolidation chemotherapy and donor lymphocyte infusion prevent second acute leukemia relapse after allotransplant	Less than 10 patients/case report
37	Cassaday R. D., Alan Potts D., Stevenson P. A., Bar M., Georges G. E., Shustov A. R., Sorror M. L., Wood B. L., Delaney C., Doney K. C., Storb R. F., Sandmaier B. M.	Evaluation of allogeneic transplantation in first or later minimal residual disease – negative remission following adult-inspired therapy for acute lymphoblastic leukemia	Less than 10 patients/case report
43	Short N. J., Jabbour E., Sasaki K., Patel K., O'Brien S. M., Cortes J. E., Garris R., Issa G. C., Garcia-Manero G., Luthra R., Thomas D., Kantarjian H., Ravandi F.	Impact of complete molecular response on survival in patients with Philadelphia chromosome-positive acute lymphoblastic leukemia	Population not of interest
49	Yoon J. H., Yhim H. Y., Kwak J. Y., Ahn J. S., Yang D. H., Lee J. J., Kim S. J., Kim J. S., Park S. J., Choi C. W., Eom H. S., Park S. K., Choi S. Y., Kim S. H., Kim D. W., Lee S.	Minimal residual disease-based effect and long-term outcome of first-line dasatinib combined with chemotherapy for adult Philadelphia chromosomepositive acute lymphoblastic leukemia	Outcomes not separable for population of interest
50	Mannis G. N., Martin T. G., Damon L. E., Andreadis C., Olin R. L., Kong K. A., Faham M., Hwang J., Ai W. Z., Gaensler K. M. L., Sayre P. H., Wolf J. L., Logan A. C.	Quantification of Acute Lymphoblastic Leukemia Clonotypes in Leukapheresed Peripheral Blood Progenitor Cells Predicts Relapse Risk after Autologous Hematopoietic Stem Cell Transplantation	Not a treatment of interest

Reference ID	Authors	Title	Reason for Exclusion
76	Kantarjian H. M., Stein A. S., Bargou R. C., Grande Garcia C., Larson R. A., Stelljes M., Gökbuget N., Zugmaier G., Benjamin J. E., Zhang A., Jia C., Topp M. S.	Blinatumomab treatment of older adults with relapsed/refractory B-precursor acute lymphoblastic leukemia: Results from two phase 2 studies	Outcomes not separable for population of interest
81	Takashima S., Miyamoto T., Kamimura T., Yoshimoto G., Yoshida S., Henzan H., Takase K., Kato K., Ito Y., Ohno Y., Nagafuji K., Eto T., Teshima T., Akashi K.	Effects of conditioning intensity in allogeneic stem cell transplantation for Philadelphia chromosome-positive acute lymphoblastic leukemia	Outcomes not separable for population of interest
83	Mo X. D., Zhang X. H., Xu L. P., Wang Y., Yan C. H., Chen H., Chen Y. H., Han W., Wang F. R., Wang J. Z., Liu K. Y., Huang X. J.	Interferon-α: A Potentially Effective Treatment for Minimal Residual Disease in Acute Leukemia/Myelodysplastic Syndrome after Allogeneic Hematopoietic Stem Cell Transplantation	Outcomes not separable for population of interest
110	Parma M., Viganò C., Fumagalli M., Colnaghi F., Colombo A., Mottadelli F., Rossi V., Elli E., Terruzzi E., Belotti A., Cazzaniga G., Pogliani E. M., Pioltelli P.	Good outcome for very high risk adult B-cell acute lymphoblastic leukaemia carrying genetic abnormalities t(4;11)(q21;q23) or t(9;22)(q34;q11), if promptly submitted to allogeneic transplantation, after obtaining a good molecular remission	Less than 10 patients/case report
118	Aldoss I., Stiller T., Cao T. M., Palmer J. M., Thomas S. H., Forman S. J., Pullarkat V.	Impact of Additional Cytogenetic Abnormalities in Adults with Philadelphia Chromosome-Positive Acute Lymphoblastic Leukemia Undergoing Allogeneic Hematopoietic Cell Transplantation	Outcomes not separable for population of interest
134	Bachanova V., Marks D. I., Zhang M. J., Wang H., De Lima M., Aljurf M. D., Arellano M., Artz A. S., Bacher U., Cahn J. Y., Chen Y. B., Copelan E. A., Drobyski W. R., Gale R. P., Greer J. P., Gupta V., Hale G. A., Kebriaei P., Lazarus H. M., Lewis I. D., Lewis V. A., Liesveld J. L., Litzow M. R., Loren A. W., Miller A. M., Norkin M., Oran B., Pidala J., Rowe J. M., Savani B. N., Saber W., Vij R., Waller E. K., Wiernik P. H., Weisdorf D. J.	Ph+ ALL patients in first complete remission have similar survival after reduced intensity and myeloablative allogeneic transplantation: Impact of tyrosine kinase inhibitor and minimal residual disease	Population not of interest
150	Helbig G., Krawczyk-Kulis M., Kopera M., Jagoda K., Rzepka P., Majewska-Tessar A., Hejla M., Kyrcz-Krzemien S.	Autologous hematopoietic stem cell transplantation for high-risk acute lymphoblastic leukemia: Non-randomized study with a maximum follow-up of more than 22 years	Population not of interest

Reference ID	Authors	Title	Reason for Exclusion
152	Bar M., Wood B. L., Radich J. P., Doney K. C., Woolfrey A. E., Delaney C., Appelbaum F. R., Gooley T. A.	Impact of minimal residual disease, detected by flow cytometry, on outcome of myeloablative hematopoietic cell transplantation for acute lymphoblastic leukemia	Outcomes not separable for population of interest
172	Appelbaum F. R.	Measurement of minimal residual disease before and after myeloablative hematopoietic cell transplantation for acute leukemia	Population not of interest
210	Gökbuget N., Kneba M., Raff T., Trautmann H., Bartram C. R., Arnold R., Fietkau R., Freund M., Ganser A., Ludwig W. D., Maschmeyer G., Rieder H., Schwartz S., Serve H., Thiel E., Brüggemann M., Hoelzer D.	Adult patients with acute lymphoblastic leukemia and molecular failure display a poor prognosis and are candidates for stem cell transplantation and targeted therapies	No outcomes of interest
211	Chen H., Liu K. Y., Xu L. P., Liu D. H., Chen Y. H., Zhao X. Y., Han W., Zhang X. H., Wang Y., Zhang Y. Y., Qin Y. Z., Liu Y. R., Huang X. J.	Administration of imatinib after allogeneic hematopoietic stem cell transplantation may improve disease-free survival for patients with Philadelphia chromosome-positive acute lymphobla stic leukemia	Outcomes not separable for population of interest
217	Mizuta S., Matsuo K., Maeda T., Yujiri T., Hatta Y., Kimura Y., Ueda Y., Kanamori H., Usui N., Akiyama H., Takada S., Yokota A., Takatsuka Y., Tamaki S., Imai K., Moriuchi Y., Miyazaki Y., Ohtake S., Ohnishi K., Naoe T.	Prognostic factors influencing clinical outcome of allogeneic hematopoietic stem cell transplantation following imatinib-based therapy in BCR-ABL-positive ALL	Outcomes not separable for population of interest
236	Gökbuget N., Basara N., Baurmann H., Beck J., Brüggemann M., Diedrich H., Güldenzoph B., Hartung G., Horst H. A., Hüttmann A., Kobbe G., Naumann R., Ratei R., Reichle A., Serve H., Stelljes M., Viardot A., Wattad M., Hoelzer D.	High single-drug activity of nelarabine in relapsed T-lymphoblastic leukemia/lymphoma offers curative option with subsequent stem cell transplantation	Outcomes not separable for population of interest
242	Doney K., Gooley T. A., Deeg H. J., Flowers M. E. D., Storb R., Appelbaum F. R.	Allogeneic Hematopoietic Cell Transplantation with Full-Intensity Conditioning for Adult Acute Lymphoblastic Leukemia: Results from a Single Center, 1998-2006	Outcomes not separable for population of interest
266	Ma X., Wu D., Sun A., Qiu H., Fu Z., Wu X., Chen S., Mohty M.	The value of monitoring minimal residual disease in the patients with donor lymphocyte infusion as intervention of relapsed/ refractory acute lymphoblastic leukemia after allogeneic hematopoietic stem cell transplantation	No outcomes of interest

Reference ID	Authors	Title	Reason for Exclusion
375	Scheuring U. J., Pfeifer H., Wassmann B., Brück P., Atta J., Petershofen E. K., Gehrke B., Gschaidmeier H., Hoelzer D., Ottmann O. G.	Early minimal residual disease (MRD) analysis during treatment of Philadelphia chromosome/Bcr-Abl-positive acute lymphoblastic leukemia with the Abl-tyrosine kinase inhibitor imatinib (STI571)	No outcomes of interest
385	Pérez-Simón J. A., Caballero D., Diez-Campelo M., Lopez-Pérez R., Mateos G., Cañizo C., Vazquez L., Vidriales B., Mateos M. V., Gonzalez M., San Miguel J. F.	Chimerism and minimal residual disease monitoring after reduced intensity conditioning (RIC) allogeneic transplantation	Outcomes not separable for population of interest
389	Yokota H., Tsuno N. H., Tanaka Y., Fukui T., Kitamura K., Hirai H., Osumi K., Itou N., Satoh H., Okabe M., Nakahara K.	Quantification of minimal residual disease in patients with e1a2 BCR-ABL-positive acute lymphoblastic leukemia using a real-time RT-PCR assay	No outcomes of interest
399	Cornish J., Oakhill A.	The management of relapsed acute lymphoblastic leukaemia	Narrative review, letter, expert opinion, etc.
442	Mitterbauer G., Nemeth P., Wacha S., Cross N. C. P., Schwarzinger I., Jaeger U., Geissler K., Greinix H. T., Kalhs P., Lechner K., Mannhalter C.	Quantification of minimal residual disease in patients with BCR-ABL- positive acute lymphoblastic leukaemia using quantitative competitive polymerase chain reaction	Population not of interest
540	Compana D., Coustan-Smith E., Janossy G.	The immunologic detection of minimal residual disease in acute leukemia	No outcomes of interest
552	Mo X. D., Zhang X. H., Xu L. P., Wang Y., Yan C. H., Chen H., Chen Y. H., Han W., Wang F. R., Wang J. Z., Liu K. Y., Huang X. J.	Comparison of outcomes after donor lymphocyte infusion with or without prior chemotherapy for minimal residual disease in acute leukemia/myelodysplastic syndrome after allogeneic hematopoietic stem cell transplantation	Outcomes not separable for population of interest
557	Milano F., Gooley T., Wood B., Woolfrey A., Flowers M. E., Doney K., Witherspoon R., Mielcarek M., Deeg J. H., Sorror M., Dahlberg A., Sandmaier B. M., Salit R., Petersdorf E., Appelbaum F. R., Delaney C.	Cord-Blood Transplantation in Patients with Minimal Residual Disease	Outcomes not separable for population of interest
573	Ravandi F., Jorgensen J. L., O'Brien S. M., Jabbour E., Thomas D. A., Borthakur G., Garris R., Huang X., Garcia-Manero G., Burger J. A., Ferrajoli A., Wierda W., Kadia T., Jain	Minimal residual disease assessed by multi-parameter flow cytometry is highly prognostic in adult patients with acute lymphoblastic leukaemia	Population not of interest

Reference ID	Authors	Title	Reason for Exclusion
	N., Wang S. A., Konoplev S., Kebriaei P., Champlin R. E., McCue D., Estrov Z., Cortes J. E., Kantarjian H. M.		
575	Zugmaier G., Gokbuget N., Klinger M., Viardot A., Stelljes M., Neumann S., Horst H. A., Marks R., Faul C., Diedrich H., Reichle A., Bruggemann M., Holland C., Schmidt M., Einsele H., Bargou R. C., Topp M. S.	Long-term survival and T-cell kinetics in relapsed/refractory ALL patients who achieved MRD response after blinatumomab treatment	Population not of interest
579	Nishiwaki S., Imai K., Mizuta S., Kanamori H., Ohashi K., Fukuda T., Onishi Y., Takahashi S., Uchida N., Eto T., Nakamae H., Yujiri T., Mori S., Nagamura-Inoue T., Suzuki R., Atsuta Y., Tanaka J.	Impact of MRD and TKI on allogeneic hematopoietic cell transplantation for Ph+ALL: a study from the adult ALL WG of the JSHCT	Outcomes not separable for population of interest
582	Ravandi F., O'Brien S. M., Cortes J. E., Thomas D. M., Garris R., Faderl S., Burger J. A., Rytting M. E., Ferrajoli A., Wierda W. G., Verstovsek S., Champlin R., Kebriaei P., McCue D. A., Huang X., Jabbour E., Garcia-Manero G., Estrov Z., Kantarjian H. M.	Long-term follow-up of a phase 2 study of chemotherapy plus dasatinib for the initial treatment of patients with Philadelphia chromosome-positive acute lymphoblastic leukemia	Population not of interest
584	Chamseddine A. N., Willekens C., De Botton S., Bourhis J. H.	Retrospective Study of Allogeneic Hematopoietic Stem Cell Transplantation in Philadelphia Chromosome-Positive Leukemia: 25 Years' Experience at Gustave Roussy Cancer Campus	Outcomes not separable for population of interest
585	Ding Z., Han M. Z., Chen S. L., Ma Q. L., Wei J. L., Pang A. M., Zhang X. Y., Liang C., Yao J. F., Cao Y. G., Feng S. Z., Jiang E. L.	Outcomes of Adults with Acute Lymphoblastic Leukemia After Autologous Hematopoietic Stem Cell Transplantation and the Significance of Pretransplantation Minimal Residual Disease: Analysis from a Single Center of China	Less than 10 patients/case report
590	Kim D. Y., Joo Y. D., Lim S. N., Kim S. D., Lee J. H., Lee J. H., Kim D. H., Kim K., Jung C. W., Kim I., Yoon S. S., Park S., Ahn J. S., Yang D. H., Lee J. J., Lee H. S., Kim Y. S., Mun Y. C., Kim H., Park J. H., Moon J. H., Sohn S. K., Lee S. M., Lee W. S., Kim K. H., Won J. H., Hyun M. S., Park J., Lee J. H., Shin H. J., Chung J. S., Lee H., Eom H. S., Lee G. W., Cho Y. U., Jang S., Park C. J., Chi H. S., Lee K. H.	Nilotinib combined with multiagent chemotherapy for newly diagnosed Philadelphia-positive acute lymphoblastic leukemia	Outcomes not separable for population of interest
593	Salek C., Folber F., Fronkova E., Prochazka B., Marinov I., Cetkovsky P., Mayer J., Doubek M.	Early MRD response as a prognostic factor in adult patients with acute lymphoblastic leukemia	Population not of interest

Reference ID	Authors	Title	Reason for Exclusion
599	Bergfelt E., Kozlowski P., Ahlberg L., Hulegardh E., Hagglund H., Karlsson K., Markuszewska-Kuczymska A., Tomaszewska-Toporska B., Smedmyr B., Astrom M., Amini R. M., Hallbook H.	Satisfactory outcome after intensive chemotherapy with pragmatic use of minimal residual disease (MRD) monitoring in older patients with Philadelphia-negative B cell precursor acute lymphoblastic leukaemia: a Swedish registry-based study	Less than 10 patients/case report
604	Kim M., Park J., Kim D. W., Kim Y. J., Jeon Y. W., Yoon J. H., Shin S. H., Yahng S. A., Lee S. E., Cho B. S., Eom K. S., Kim H. J., Min C. K., Cho S. G., Kim Y., Lee J. W., Han K., Min W. S., Lee S.	Impact of IKZF1 deletions on long-term outcomes of allo-SCT following imatinib-based chemotherapy in adult Philadelphia chromosome-positive ALL	Population not of interest
606	Kong Y., Xu L. P., Liu Y. R., Qin Y. Z., Sun Y. Q., Wang Y., Jiang H., Jiang Q., Chen H., Chang Y. J., Huang X. J.	Presence of CD34(+)CD38(-)CD58(-) leukemia-propagating cells at diagnosis identifies patients at high risk of relapse with Ph chromosome-positive ALL after allo-hematopoietic SCT	Population not of interest
608	Topp M. S., Gokbuget N., Zugmaier G., Klappers P., Stelljes M., Neumann S., Viardot A., Marks R., Diedrich H., Faul C., Reichle A., Horst H. A., Bruggemann M., Wessiepe D., Holland C., Alekar S., Mergen N., Einsele H., Hoelzer D., Bargou R. C.	Phase II trial of the anti-CD19 bispecific T cell-engager blinatumomab shows hematologic and molecular remissions in patients with relapsed or refractory B-precursor acute lymphoblastic leukemia	Population not of interest
609	Kanakry C. G., Tsai H. L., Bolanos-Meade J., Smith B. D., Gojo I., Kanakry J. A., Kasamon Y. L., Gladstone D. E., Matsui W., Borrello I., Huff C. A., Swinnen L. J., Powell J. D., Pratz K. W., DeZern A. E., Showel M. M., McDevitt M. A., Brodsky R. A., Levis M. J., Ambinder R. F., Fuchs E. J., Rosner G. L., Jones R. J., Luznik L.	Single-agent GVHD prophylaxis with posttransplantation cyclophosphamide after myeloablative, HLA-matched BMT for AML, ALL, and MDS	Population not of interest
611	Ma L., Hao S., Diong C., Goh Y. T., Gopalakrishnan S., Ho A., Hwang W., Koh L. P., Koh M., Lim Z. Y., Loh Y., Poon M., Tan L. K., Tan P., Linn Y. C.	Pre-transplant achievement of negativity in minimal residual disease and French-American-British L1 morphology predict superior outcome after allogeneic transplant for Philadelphia chromosome positive acute lymphoblastic leukemia: an analysis of Southeast Asian patients	Outcomes not separable for population of interest
618	Tucunduva L., Ruggeri A., Sanz G., Furst S., Cornelissen J., Linkesch W., Mannone L., Ribera J. M., Veelken H., Yakoub-	Impact of minimal residual disease on outcomes after umbilical cord blood transplantation for adults with Philadelphia-positive acute	Not a treatment of interest

Reference ID	Authors	Title	Reason for Exclusion
	Agha I., Gonzalez Valentin M. E., Schots R., Arcese W., Montesinos P., Labopin M., Gluckman E., Mohty M., Rocha V.	lymphoblastic leukaemia: an analysis on behalf of Eurocord, Cord Blood Committee and the Acute Leukaemia working party of the European group for Blood and Marrow Transplantation	
620	Terwey T. H., Hemmati P. G., Nagy M., Pfeifer H., Gokbuget N., Bruggemann M., Le Duc T. M., le Coutre P., Dorken B., Arnold R.	Comparison of chimerism and minimal residual disease monitoring for relapse prediction after allogeneic stem cell transplantation for adult acute lymphoblastic leukemia	Population not of interest
622	Logan A. C., Vashi N., Faham M., Carlton V., Kong K., Buno I., Zheng J., Moorhead M., Klinger M., Zhang B., Waqar A., Zehnder J. L., Miklos D. B.	Immunoglobulin and T cell receptor gene high-throughput sequencing quantifies minimal residual disease in acute lymphoblastic leukemia and predicts post-transplantation relapse and survival	Population not of interest
623	Yan C. H., Jiang Q., Wang J., Xu L. P., Liu D. H., Jiang H., Chen H., Zhang X. H., Liu K. Y., Huang X. J.	Superior survival of unmanipulated haploidentical hematopoietic stem cell transplantation compared with chemotherapy alone used as post-remission therapy in adults with standard-risk acute lymphoblastic leukemia in first complete remission	Outcomes not separable for population of interest
625	Salah-Eldin M., Abousamra N. K., Azzam H.	Clinical significance of minimal residual disease in young adults with standard-risk/Ph-negative precursor B-acute lymphoblastic leukemia: results of prospective study	Population not of interest
627	Zhou Y., Slack R., Jorgensen J. L., Wang S. A., Rondon G., de Lima M., Shpall E., Popat U., Ciurea S., Alousi A., Qazilbash M., Hosing C., O'Brien S., Thomas D., Kantarjian H., Medeiros L. J., Champlin R. E., Kebriaei P.	The effect of peritransplant minimal residual disease in adults with acute lymphoblastic leukemia undergoing allogeneic hematopoietic stem cell transplantation	Population not of interest
629	Tan Y., Du K., Luo Y., Shi J., Cao L., Zheng Y., Zheng G., Zhao Y., Ye X., Cai Z., Huang H.	Superiority of preemptive donor lymphocyte infusion based on minimal residual disease in acute leukemia patients after allogeneic hematopoietic stem cell transplantation	Outcomes not separable for population of interest
632	Gossai N., Verneris M. R., Karras N. A., Gorman M. F., Patel N. J., Burke M. J.	A clofarabine-based bridging regimen in patients with relapsed ALL and persistent minimal residual disease (MRD)	Less than 10 patients/case report

Reference ID	Authors	Title	Reason for Exclusion
634	Giebel S., Labopin M., Gorin N. C., Caillot D., Leguay T., Schaap N., Michallet M., Dombret H., Mohty M.	Improving results of autologous stem cell transplantation for Philadelphia-positive acute lymphoblastic leukaemia in the era of tyrosine kinase inhibitors: a report from the Acute Leukaemia Working Party of the European Group for Blood and Marrow Transplantation	Less than 10 patients/case report
637	Kunter G., Perkins J. B., Pidala J., Nishihori T., Kharfan-Dabaja M. A., Field T., Fernandez H., Perez L., Locke F., Ayala E., Tomblyn M., Ochoa-Bayona J. L., Betts B., Nieder M., Anasetti C.	Pharmacokinetically-targeted BU and fludarabine as conditioning before allogeneic hematopoietic cell transplantation for adults with ALL in first remission	Not a treatment of interest
640	Samra M. A., Mahmoud H. K., Abdelhamid T. M., El Sharkawy N. M., Elnahass Y. H., Elgammal M., Abdelfattah R. M., Eid S., Ghaleb F. M., Kamel A. M.	The prognostic significance of minimal residual disease in adult Egyptian patients with precursor acute lymphoblastic leukemia	Population not of interest
641	Yan C. H., Wang J. Z., Liu D. H., Xu L. P., Chen H., Liu K. Y., Huang X. J.	Chemotherapy followed by modified donor lymphocyte infusion as a treatment for relapsed acute leukemia after haploidentical hematopoietic stem cell transplantation without in vitro T-cell depletion: superior outcomes compared with chemotherapy alone and an analysis of prognostic factors	Outcomes not separable for population of interest
642	Ravandi F., Jorgensen J. L., Thomas D. A., O'Brien S., Garris R., Faderl S., Huang X., Wen S., Burger J. A., Ferrajoli A., Kebriaei P., Champlin R. E., Estrov Z., Challagundla P., Wang S. A., Luthra R., Cortes J. E., Kantarjian H. M.	Detection of MRD may predict the outcome of patients with Philadelphia chromosome-positive ALL treated with tyrosine kinase inhibitors plus chemotherapy	Population not of interest
653	Nagafuji K., Miyamoto T., Eto T., Kamimura T., Taniguchi S., Okamura T., Ohtsuka E., Yoshida T., Higuchi M., Yoshimoto G., Fujisaki T., Abe Y., Takamatsu Y., Yokota S., Akashi K., Harada M.	Monitoring of minimal residual disease (MRD) is useful to predict prognosis of adult patients with Ph-negative ALL: results of a prospective study (ALL MRD2002 Study)	Population not of interest
654	Kebriaei P., Wilhelm K., Ravandi F., Brandt M., de Lima M., Ciurea S., Worth L., O'Brien S., Thomas D., Champlin R. E., Kantarjian H.	Feasibility of allografting in patients with advanced acute lymphoblastic leukemia after salvage therapy with inotuzumab ozogamicin	Outcomes not separable for population of interest
657	Pfeifer H., Wassmann B., Bethge W., Dengler J., Bornhauser M., Stadler M., Beelen D., Vucinic V., Burmeister T., Stelljes M., Faul C., Dreger P., Kiani A., Schafer-Eckart K., Schwerdtfeger	Randomized comparison of prophylactic and minimal residual disease-triggered imatinib after	Population not of interest

Reference ID	Authors	Title	Reason for Exclusion
	R., Lange E., Kubuschok B., Horst H. A., Gramatzki M., Bruck P., Serve H., Hoelzer D., Gokbuget N., Ottmann O. G.	allogeneic stem cell transplantation for BCR-ABL1-positive acute lymphoblastic leukemia	
666	Lee S., Kim D. W., Cho B. S., Yoon J. H., Shin S. H., Yahng S. A., Lee S. E., Eom K. S., Kim Y. J., Chung N. G., Kim H. J., Min C. K., Lee J. W., Min W. S., Park C. W.	Impact of minimal residual disease kinetics during imatinib-based treatment on transplantation outcome in Philadelphia chromosome-positive acute lymphoblastic leukemia	No outcomes of interest
693	Kikuchi M., Tanaka J., Kondo T., Hashino S., Kasai M., Kurosawa M., Iwasaki H., Morioka M., Kawamura T., Masauzi N., Fukuhara T., Kakinoki Y., Kobayashi H., Noto S., Asaka M., Imamura M.	Clinical significance of minimal residual disease in adult acute lymphoblastic leukemia	No outcomes of interest
695	Nishiwaki S., Miyamura K., Kato C., Terakura S., Ohashi K., Sakamaki H., Nakao S., Harigae H., Kodera Y.	Impact of post-transplant imatinib administration on Philadelphia chromosome-positive acute lymphoblastic leukaemia	Outcomes not separable for population of interest
702	Patel B., Rai L., Buck G., Richards S. M., Mortuza Y., Mitchell W., Gerrard G., Moorman A. V., Duke V., Hoffbrand A. V., Fielding A. K., Goldstone A. H., Foroni L.	Minimal residual disease is a significant predictor of treatment failure in non T-lineage adult acute lymphoblastic leukaemia: final results of the international trial UKALL XII/ECOG2993	Population not of interest
703	Giebel S., Stella-Holowiecka B., Krawczyk-Kulis M., Gokbuget N., Hoelzer D., Doubek M., Mayer J., Piatkowska-Jakubas B., Skotnicki A. B., Dombret H., Ribera J. M., Piccaluga P. P., Czerw T., Kyrcz-Krzemien S., Holowiecki J.	Status of minimal residual disease determines outcome of autologous hematopoietic SCT in adult ALL	Population not of interest
711	Gutman J. A., Leisenring W., Appelbaum F. R., Woolfrey A. E., Delaney C.	Low relapse without excessive transplant-related mortality following myeloablative cord blood transplantation for acute leukemia in complete remission: a matched cohort analysis	Outcomes not separable for population of interest
716	Bassan R., Spinelli O., Oldani E., Intermesoli T., Tosi M., Peruta B., Rossi G., Borlenghi E., Pogliani E. M., Terruzzi E., Fabris P., Cassibba V., Lambertenghi-Deliliers G., Cortelezzi A., Bosi A., Gianfaldoni G., Ciceri F., Bernardi M., Gallamini A., Mattei D., Di Bona E., Romani C., Scattolin A. M., Barbui T., Rambaldi A.	Improved risk classification for risk-specific therapy based on the molecular study of minimal residual disease (MRD) in adult acute lymphoblastic leukemia (ALL)	Population not of interest
721	Yanada M., Sugiura I., Takeuchi J., Akiyama H., Maruta A., Ueda Y., Usui N., Yagasaki F., Yujiri T., Takeuchi M., Nishii K.,	Prospective monitoring of BCR-ABL1 transcript levels in patients with Philadelphia chromosome-	No outcomes of interest

Reference ID	Authors	Title	Reason for Exclusion
	Kimura Y., Miyawaki S., Narimatsu H., Miyazaki Y., Ohtake S., Jinnai I., Matsuo K., Naoe T., Ohno R.	positive acute lymphoblastic leukaemia undergoing imatinib-combined chemotherapy	
724	Holowiecki J., Krawczyk-Kulis M., Giebel S., Jagoda K., Stella-Holowiecka B., Piatkowska-Jakubas B., Paluszewska M., Seferynska I., Lewandowski K., Kielbinski M., Czyz A., Balana-Nowak A., Krol M., Skotnicki A. B., Jedrzejczak W. W., Warzocha K., Lange A., Hellmann A.	Status of minimal residual disease after induction predicts outcome in both standard and high-risk Ph-negative adult acute lymphoblastic leukaemia. The Polish Adult Leukemia Group ALL 4-2002 MRD Study	Population not of interest
736	Spinelli O., Peruta B., Tosi M., Guerini V., Salvi A., Zanotti M. C., Oldani E., Grassi A., Intermesoli T., Mico C., Rossi G., Fabris P., Lambertenghi-Deliliers G., Angelucci E., Barbui T., Bassan R., Rambaldi A.	Clearance of minimal residual disease after allogeneic stem cell transplantation and the prediction of the clinical outcome of adult patients with high-risk acute lymphoblastic leukemia	Population not of interest
742	Raff T., Gokbuget N., Luschen S., Reutzel R., Ritgen M., Irmer S., Bottcher S., Horst H. A., Kneba M., Hoelzer D., Bruggemann M.	Molecular relapse in adult standard-risk ALL patients detected by prospective MRD monitoring during and after maintenance treatment: data from the GMALL 06/99 and 07/03 trials	Population not of interest
748	Toubai T., Tanaka J., Ota S., Fukuhara T., Hashino S., Kondo T., Kasai M., Kakinoki Y., Masauzi N., Morioka M., Kawamura T., Iwasaki H., Asaka M., Imamura M.	Minimal residual disease (MRD) monitoring using rearrangement of T-cell receptor and immunoglobulin H gene in the treatment of adult acute lymphoblastic leukemia patients	Less than 10 patients/case report
749	Bruggemann M., Raff T., Flohr T., Gokbuget N., Nakao M., Droese J., Luschen S., Pott C., Ritgen M., Scheuring U., Horst H. A., Thiel E., Hoelzer D., Bartram C. R., Kneba M.	Clinical significance of minimal residual disease quantification in adult patients with standard-risk acute lymphoblastic leukemia	Population not of interest
752	Wassmann B., Pfeifer H., Stadler M., Bornhauser M., Bug G., Scheuring U. J., Bruck P., Stelljes M., Schwerdtfeger R., Basara N., Perz J., Bunjes D., Ledderose G., Mahlberg R., Binckebanck A., Gschaidmeier H., Hoelzer D., Ottmann O. G.	Early molecular response to posttransplantation imatinib determines outcome in MRD+ Philadelphia-positive acute lymphoblastic leukemia (Ph+ ALL)	Population not of interest
773	Scheuring U. J., Pfeifer H., Wassmann B., Bruck P., Gehrke B., Petershofen E. K., Gschaidmeier H., Hoelzer D., Ottmann O. G.	Serial minimal residual disease (MRD) analysis as a predictor of response duration in Philadelphia-positive acute lymphoblastic leukemia (Ph+ALL) during imatinib treatment	Less than 10 patients/case report
777	Lee S., Kim D. W., Kim Y. J., Chung N. G., Kim Y. L., Hwang J. Y., Kim C. C.	Minimal residual disease-based role of imatinib as a first-line interim therapy prior to allogeneic stem	No outcomes of interest

Reference ID	Authors	Title	Reason for Exclusion
		cell transplantation in Philadelphia chromosome-positive acute lymphoblastic leukemia	
779	Esperou H., Boiron J. M., Cayuela J. M., Blanchet O., Kuentz M., Jouet J. P., Milpied N., Cahn J. Y., Faucher C., Bourhis J. H., Michallet M., Tanguy M. L., Vernant J. P., Gabert J., Bordigoni P., Ifrah N., Baruchel A., Dombret H.	A potential graft-versus-leukemia effect after allogeneic hematopoietic stem cell transplantation for patients with Philadelphia chromosome-positive acute lymphoblastic leukemia: results from the French Bone Marrow Transplantation Society	Outcomes not separable for population of interest
780	Stirewalt D. L., Guthrie K. A., Beppu L., Bryant E. M., Doney K., Gooley T., Appelbaum F. R., Radich J. P.	Predictors of relapse and overall survival in Philadelphia chromosome-positive acute lymphoblastic leukemia after transplantation	Outcomes not separable for population of interest
781	Szatrowski T. P., Dodge R. K., Reynolds C., Westbrook C. A., Frankel S. R., Sklar J., Stewart C. C., Hurd D. D., Kolitz J. E., Velez-Garcia E., Stone R. M., Bloomfield C. D., Schiffer C. A., Larson R. A.	Lineage specific treatment of adult patients with acute lymphoblastic leukemia in first remission with anti-B4-blocked ricin or high-dose cytarabine: Cancer and Leukemia Group B Study 9311	No outcomes of interest
789	Miglino M., Berisso G., Grasso R., Canepa L., Clavio M., Pierri I., Pietrasanta D., Gatto S., Varaldo R., Ballerini F., Verdiani S., Casarino L., DeStefano F., Sessarego M., Dominietto A., Raiola A. M., Bregante S., di Grazia C., Gobbi M., Bacigalupo A.	Allogeneic bone marrow transplantation (BMT) for adults with acute lymphoblastic leukemia (ALL): predictive role of minimal residual disease monitoring on relapse	Population not of interest
794	Dombret H., Gabert J., Boiron J. M., Rigal-Huguet F., Blaise D., Thomas X., Delannoy A., Buzyn A., Bilhou-Nabera C., Cayuela J. M., Fenaux P., Bourhis J. H., Fegueux N., Charrin C., Boucheix C., Lheritier V., Esperou H., MacIntyre E., Vernant J. P., Fiere D.	Outcome of treatment in adults with Philadelphia chromosome-positive acute lymphoblastic leukemia--results of the prospective multicenter LALA-94 trial	Population not of interest
798	Sanchez J., Serrano J., Gomez P., Martinez F., Martin C., Madero L., Herrera C., Garcia J. M., Casano J., Torres A.	Clinical value of immunological monitoring of minimal residual disease in acute lymphoblastic leukaemia after allogeneic transplantation	Less than 10 patients/case report
804	Uzunel M., Mattsson J., Jaksch M., Remberger M., Ringden O.	The significance of graft-versus-host disease and pretransplantation minimal residual disease status to outcome after allogeneic stem cell transplantation in patients with acute lymphoblastic leukemia	Outcomes not separable for population of interest

Reference ID	Authors	Title	Reason for Exclusion
810	Malec M., Bjorklund E., Soderhall S., Mazur J., Sjogren A. M., Pisa P., Bjorkholm M., Porwit-MacDonald A.	Flow cytometry and allele-specific oligonucleotide PCR are equally effective in detection of minimal residual disease in ALL	No outcomes of interest
812	Wakasugi S., Ohta K., Hasegawa Y., Tatumi N., Nakamura H.	Detection of minimal residual disease in acute leukemia by Tc-99m MIBI femoral marrow imaging	Population not of interest
819	Finke J., Bertz H., Schmoor C., Veelken H., Behringer D., Wasch R., Kunzmann R., Heidecker L., Lang H., Meyer-Konig U., Mertelsmann R.	Allogeneic bone marrow transplantation from unrelated donors using in vivo anti-T-cell globulin	Outcomes not separable for population of interest
825	Munoz L., Lopez O., Martino R., Brunet S., Bellido M., Rubiol E., Sierra J., Nomdedeu J. F.	Combined use of reverse transcriptase polymerase chain reaction and flow cytometry to study minimal residual disease in Philadelphia positive acute lymphoblastic leukemia	No outcomes of interest
836	Cimino G., Elia L., Rapanotti M. C., Sprovieri T., Mancini M., Cuneo A., Mecucci C., Fioritoni G., Carotenuto M., Morra E., Liso V., Annino L., Saglio G., De Rossi G., Foa R., Mandelli F.	A prospective study of residual-disease monitoring of the ALL1/AF4 transcript in patients with t(4;11) acute lymphoblastic leukemia	No outcomes of interest
881	Radich J., Ladne P., Gooley T.	Polymerase chain reaction-based detection of minimal residual disease in acute lymphoblastic leukemia predicts relapse after allogeneic BMT	Population not of interest
908	Uckun F. M., Kersey J. H., Haake R., Weisdorf D., Nesbit M. E., Ramsay N. K.	Pretransplantation burden of leukemic progenitor cells as a predictor of relapse after bone marrow transplantation for acute lymphoblastic leukemia	Outcomes not separable for population of interest
909	Nizet Y., Van Daele S., Lewalle P., Vaerman J. L., Philippe M., Vermylen C., Cornu G., Ferrant A., Michaux J. L., Martiat P.	Long-term follow-up of residual disease in acute lymphoblastic leukemia patients in complete remission using clonogeneic IgH probes and the polymerase chain reaction	Less than 10 patients/case report
917	Miyamura K., Tanimoto M., Morishima Y., Horibe K., Yamamoto K., Akatsuka M., Kodera Y., Kojima S., Matsuyama K., Hirabayashi N., et al.,	Detection of Philadelphia chromosome-positive acute lymphoblastic leukemia by polymerase chain reaction: possible eradication of minimal residual disease by marrow transplantation	Less than 10 patients/case report
918	Uckun F. M., Kersey J. H., Haake R., Weisdorf D., Ramsay N. K.	Autologous bone marrow transplantation in high-risk remission B-lineage acute lymphoblastic leukemia using a cocktail of three monoclonal	Outcomes not separable for population of interest

Reference ID	Authors	Title	Reason for Exclusion
		antibodies (BA-1/CD24, BA-2/CD9, and BA-3/CD10) plus complement and 4-hydroperoxycyclophosphamide for ex vivo bone marrow purging	
926	Gehly G. B., Bryant E. M., Lee A. M., Kidd P. G., Thomas E. D.	Chimeric BCR-abl messenger RNA as a marker for minimal residual disease in patients transplanted for Philadelphia chromosome-positive acute lymphoblastic leukemia	No outcomes of interest
933	Uckun F. M., Kersey J. H., Vallera D. A., Ledbetter J. A., Weisdorf D., Myers D. E., Haake R., Ramsay N. K.	Autologous bone marrow transplantation in high-risk remission T-lineage acute lymphoblastic leukemia using immunotoxins plus 4-hydroperoxycyclophosphamide for marrow purging	Less than 10 patients/case report
934	Zhang R. X., Yao E. G.	Significance of FCM-DNA measurement in detecting minimal residual disease in leukemia	Population not of interest
936	Hoelzer D.	Change in treatment strategies for adult acute lymphoblastic leukemia (ALL) according to prognostic factors and minimal residual disease	Narrative review, letter, expert opinion, etc.
938	Campana D., Coustan-Smith E., Janossy G.	The immunologic detection of minimal residual disease in acute leukemia	Outcomes not separable for population of interest
940	Rizzoli V., Mangoni L.	Pharmacological-mediated purging with mafosfamide in acute and chronic myeloid leukemias. The Italian Study Group	Population not of interest
967	May Mb, Glode A	Blinatumomab: A novel, bispecific, T-cell engaging antibody	Narrative review, letter, expert opinion, etc.
968	DeAngelo Dj, Stelljes M, Martinelli G, Kantarjian H, Liedtke M, Stock W, Goekbuget N, Wang K, Pacagnella L, Sleight B, Vandendries E, Advani As	Efficacy and safety of inotuzumab ozogamicin (INO) vs standard of care (SoC) in salvage 1 or 2 patients with acute lymphoblastic leukemia (ALL): An ongoing global phase 3 study	No outcomes of interest

Reference ID	Authors	Title	Reason for Exclusion
971	Bertrand Y, Baruchel A, Thomas Xg, Blin N, Tardy Et, Perel Y, Vey N, Gandemer V, Cacheux V, Mazingue F, Raffoux E, Plat G, Poiree M, Stephan J-L, Auvrignon A, Plantaz D, Pellier I, Bonin C, El-Hariry I, Ferster A	Clinical activity of ERY001 (erythrocyte encapsulated l-asparaginase) and native l-asparaginase (L-ASP) in combination with COOPRALL regimen in phase III randomized trial in patients with relapsed acute lymphoblastic leukemia (ALL)	Narrative review, letter, expert opinion, etc.
981	Mo X-D, Zhang X-H, Xu L-P, Wang Y, Yan C-H, Chen H, Chen Y-H, Han W, Wang F-R, Wang J-Z, Liu K-Y, Huang X-J	Comparison of outcomes after donor lymphocyte infusion with or without prior chemotherapy for minimal residual disease in acute leukemia/myelodysplastic syndrome after allogeneic hematopoietic stem cell transplantation	Outcomes not separable for population of interest

Appendix N MRD Status epidemiology estimates

Weighted average estimates of MRD+ status in Ph-negative ALL patients derived from 3 publications are provided below in Table 88.

Table 88. Weighted average of MRD+ in Ph-negative ALL patients

Study	Patients	Patients in CR	Patients in MolCR tested after induction	Patients in MolCR	% MolCR	Weights	Weighted % MolCR
Gokbuget et al. (2012)*	1,076	961	383	252	0.66	0.665	0.438
Holowiecki et al. (2008)†	131	118	115	77	0.67	0.200	0.134
Bassan et al. (2009)‡	148	125	78	41	0.53	0.135	0.071
Total patients in MolCR	64.2%						
Total weighted average MRD+ (%)	36%						

Footnotes: *B-lineage ALL patients, MRD detection threshold of 10^{-4}; †MRD detection threshold of 10^{-3}; ‡B-lineage Ph-positive ALL patients, MRD detection limit of 10^{-4}.
Abbreviations: MRD: minimal residual disease; Ph: Philadelphia chromosome; ALL: acute lymphoblastic leukaemia; CR: complete response; MolCR: molecular complete response (i.e. MRD-).
Sources: Gökbuget et al. (2012),[116] Holowiecki et al. (2008),[31] Bassan et al. (2009)[25]

Appendix O Change in MRD status after cycle 1

Table 89. Change in MRD result from baseline to end of cycle 1 in MRD non-responders

	Cycle 1 MRD[2] Value						
Baseline MRD[1]	**-5**	**-4**	**-3**	**-2**	**-1**	**NA**	**Total**
NA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
-5	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
-4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
-3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
-2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
-1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: [1]If multiple MRD measurements were taken at baseline visit, the latest one prior to blinatumomab infusion was used. [2]If multiple MRD measurements were taken during cycle 1, the non-missing minimum MRD result was used.
Note: MRD units are in 10^{-x} (i.e., -5 = 10^{-5}, etc)
Abbreviations: MRD: minimal residual disease.
Source: BLAST Key Secondary Analysis CSR Table 14-4.5.2[56]

Appendix P: RFS and OS distributions fitted to IPTW-ATT Propensity Matched Data from BLAST and Historical Control

1.1 Relapse-Free and Overall Survival Using ATT Weights

1.1.1 Relapse-Free Survival

Figure 57. Kaplan Meier estimates of RFS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATT weights

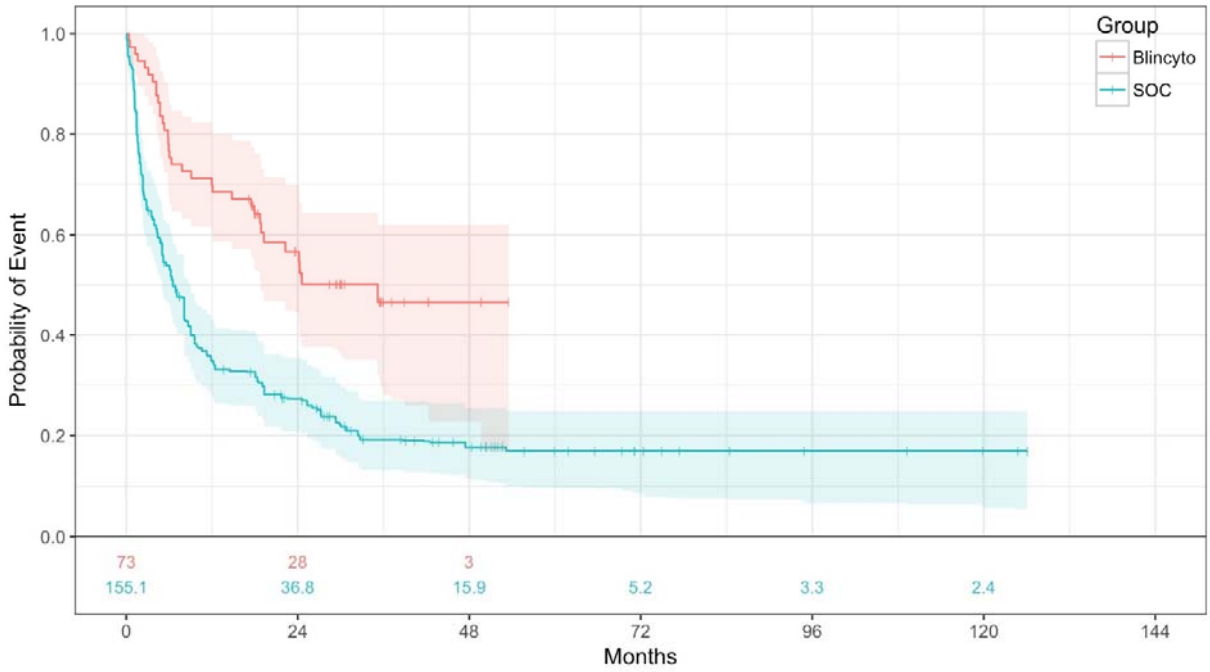


Figure 58. Fit statistics for parametric survival distributions fit to RFS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATT weights

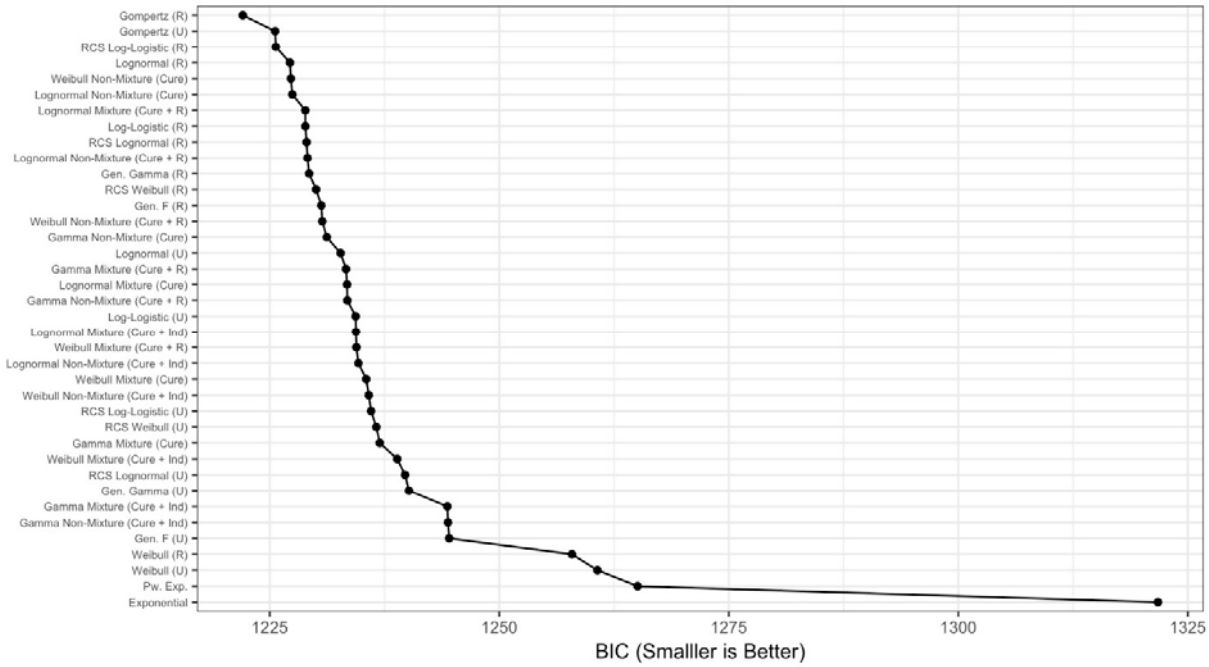


Figure 59. Treatment effect counterfactual plots for RFS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATT weights

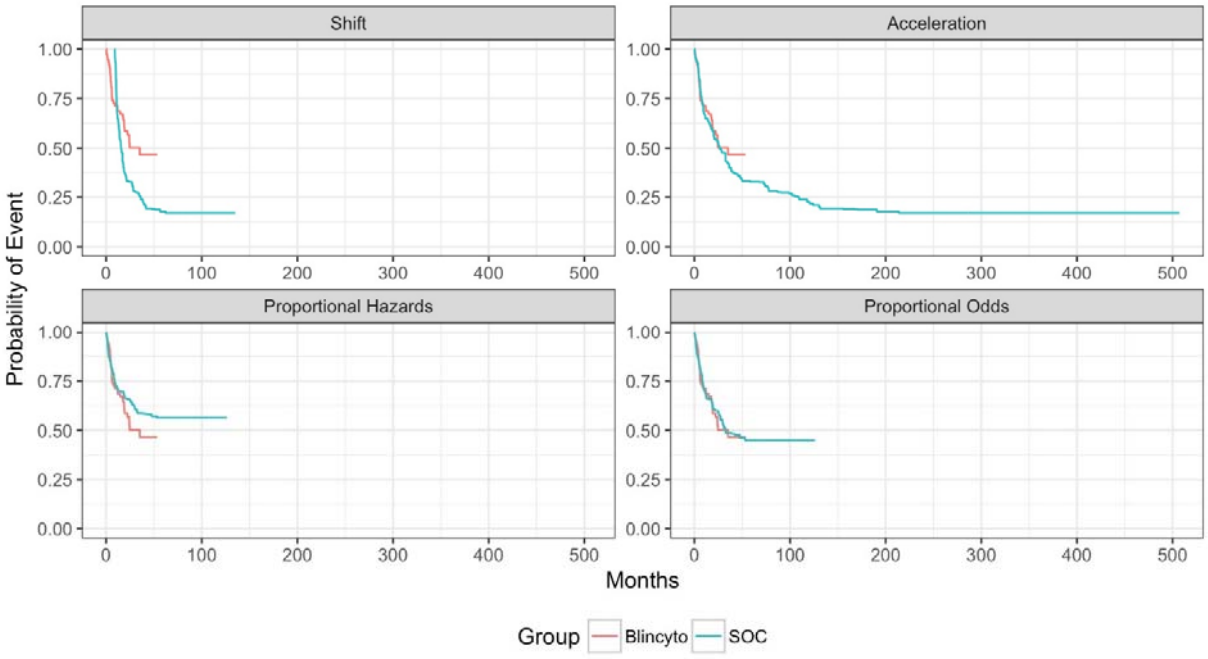


Figure 60. Schoenfeld residual plots for RFS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATT weights

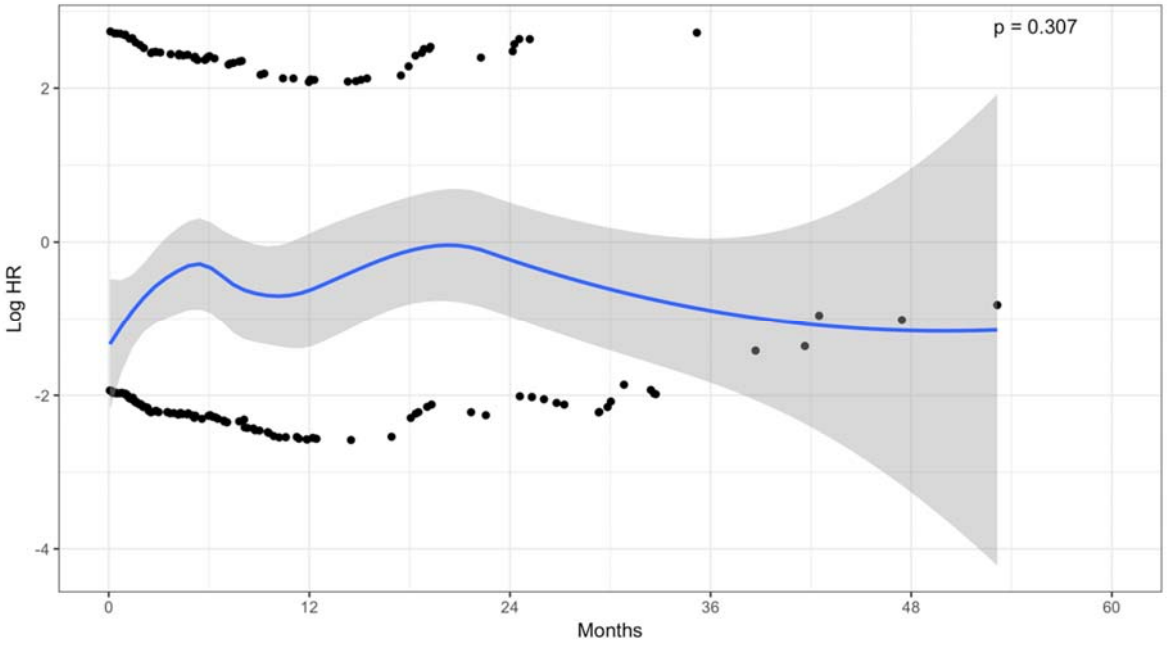
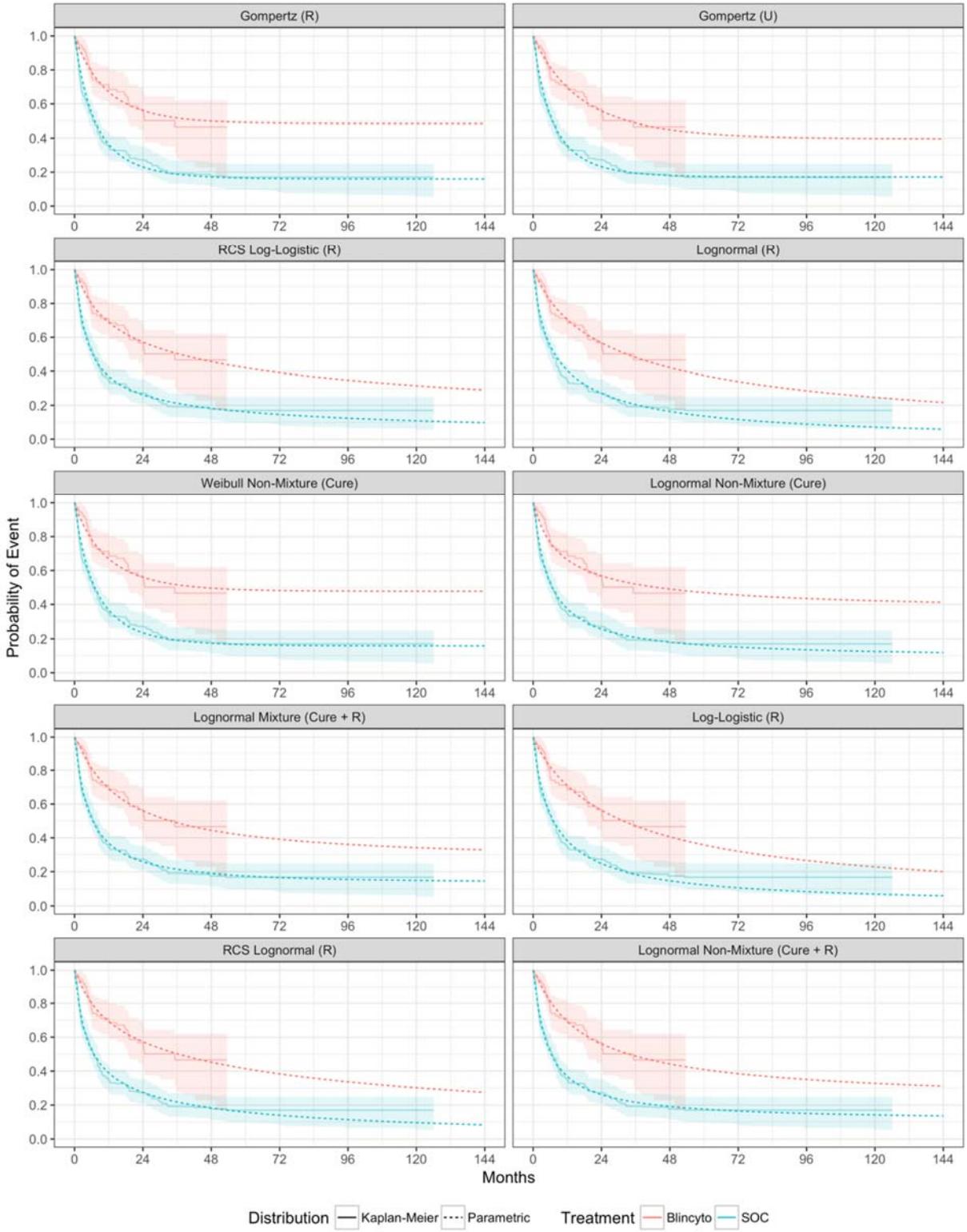
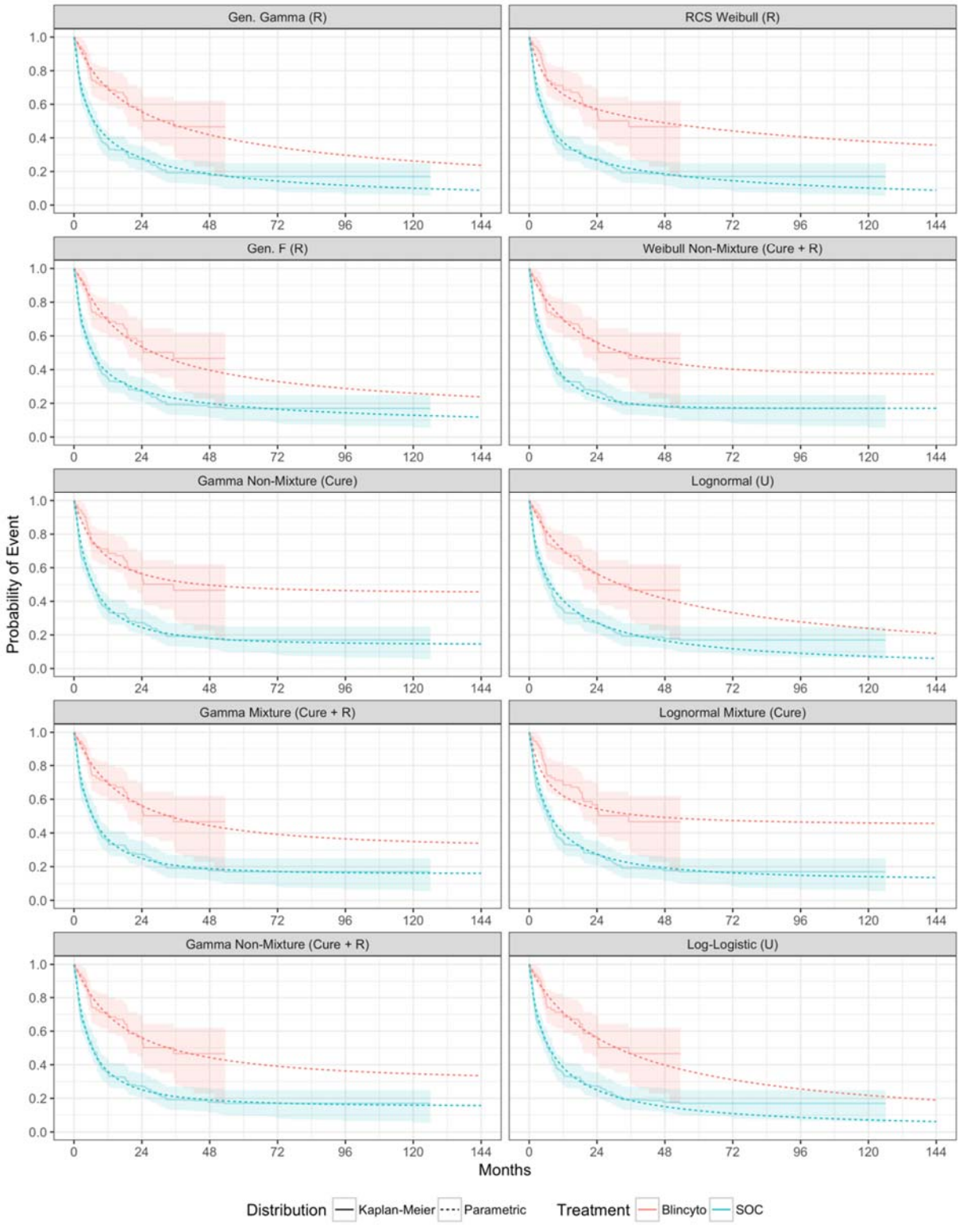
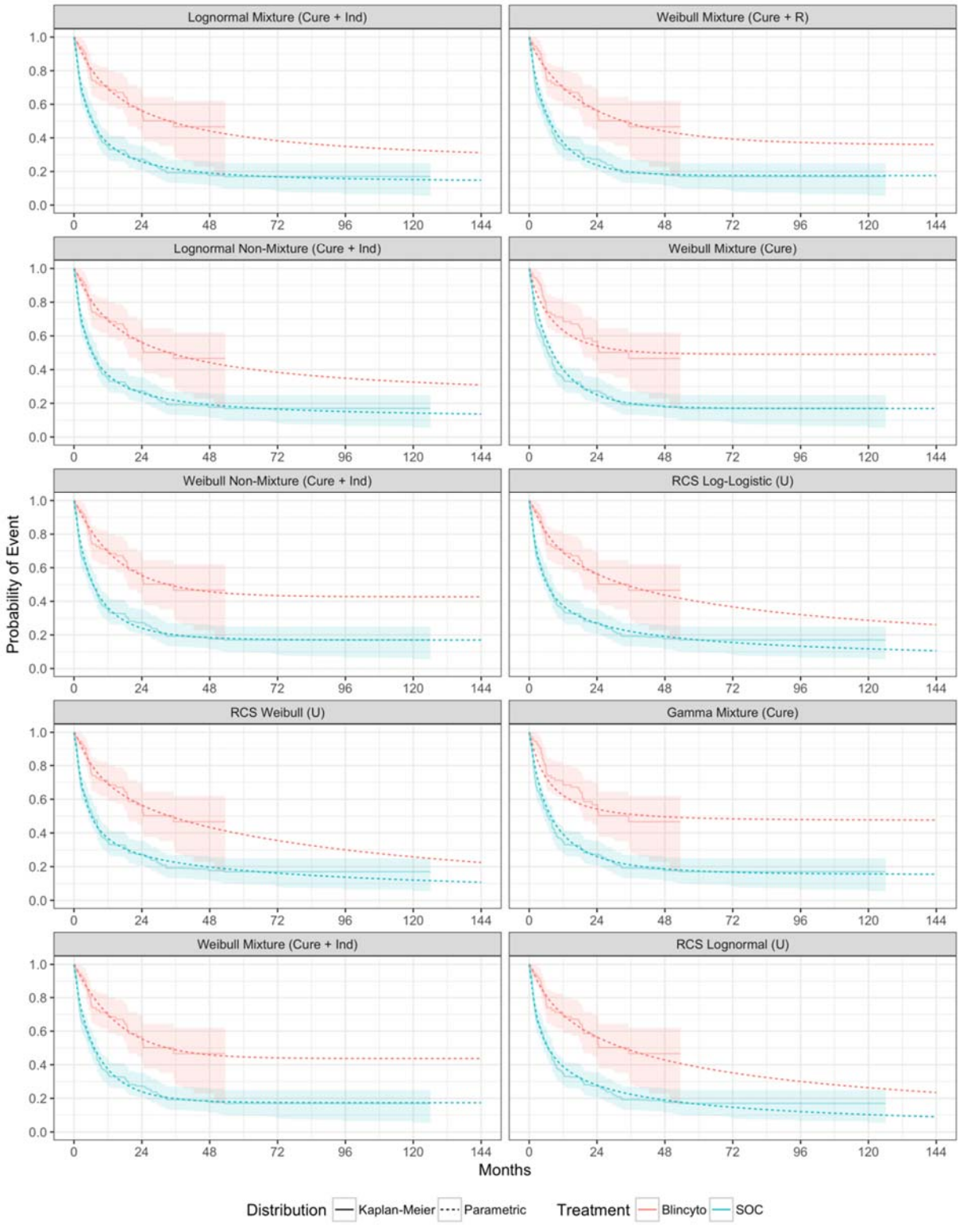


Figure 61. Survival probabilities to 12 years for all parametric survival distributions fit to RFS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATT weights




Gen. Gamma (R)
RCS Weibull (R)
Gen. F (R)
Weibull Non-Mixture (Cure + R)
Gamma Non-Mixture (Cure)
Lognormal (U)
Gamma Mixture (Cure + R)
Lognormal Mixture (Cure)
Gamma Non-Mixture (Cure + R)
Log-Logistic (U)
Probability of Event
0.0
0.2
0.4
0.6
0.8
1.0
0
24
48
72
96
120
144
Months
Distribution
Kaplan-Meier
Parametric
Treatment
Blincyto
SOC


Lognormal Mixture (Cure + Ind)
Weibull Mixture (Cure + R)
Lognormal Non-Mixture (Cure + Ind)
Weibull Mixture (Cure)
Weibull Non-Mixture (Cure + Ind)
RCS Log-Logistic (U)
RCS Weibull (U)
Gamma Mixture (Cure)
Weibull Mixture (Cure + Ind)
RCS Lognormal (U)
Probability of Event
0.0
0.2
0.4
0.6
0.8
1.0
0
24
48
72
96
120
144
Months
Distribution
Kaplan-Meier
Parametric
Treatment
Blincyto
SOC

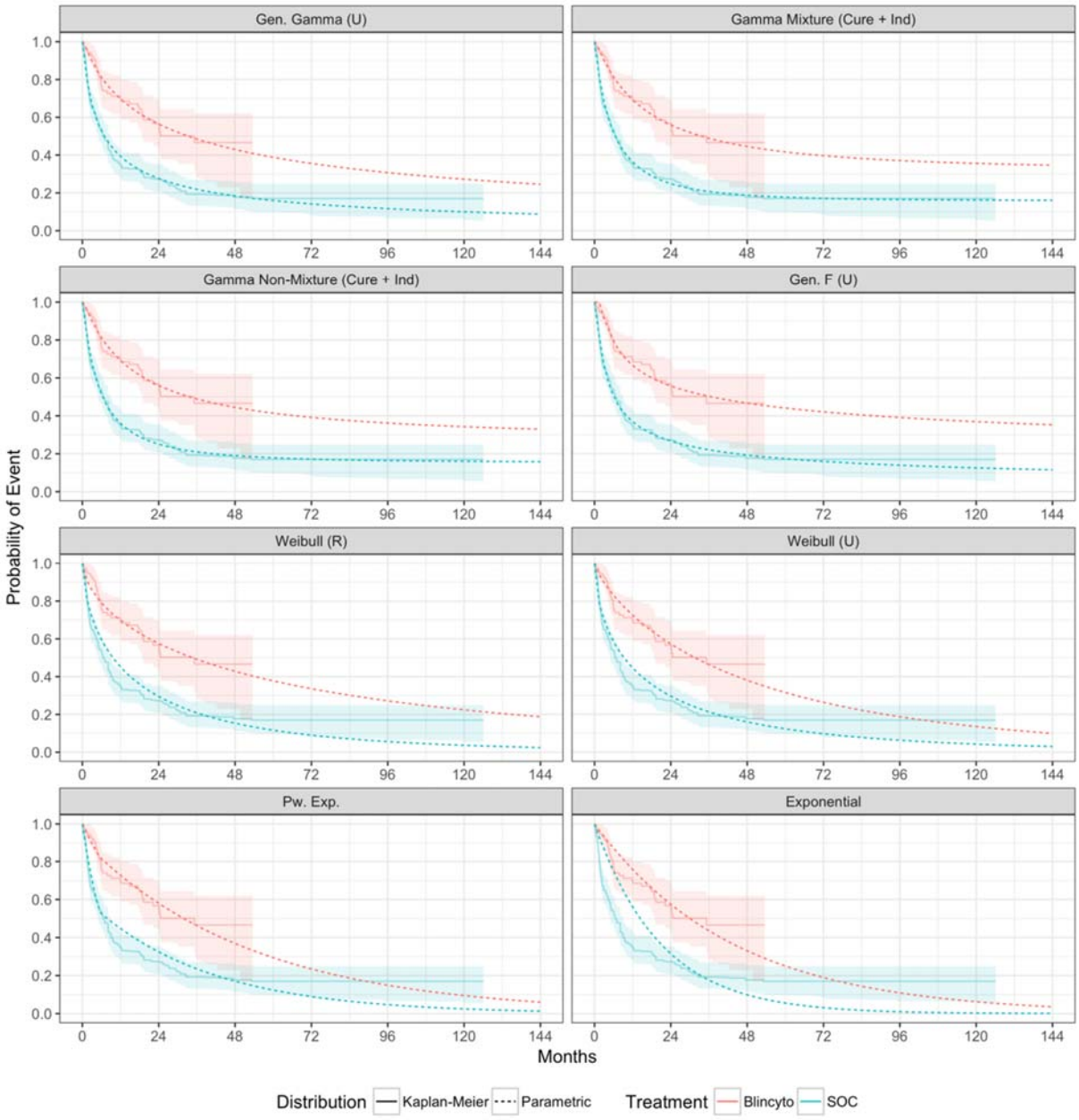
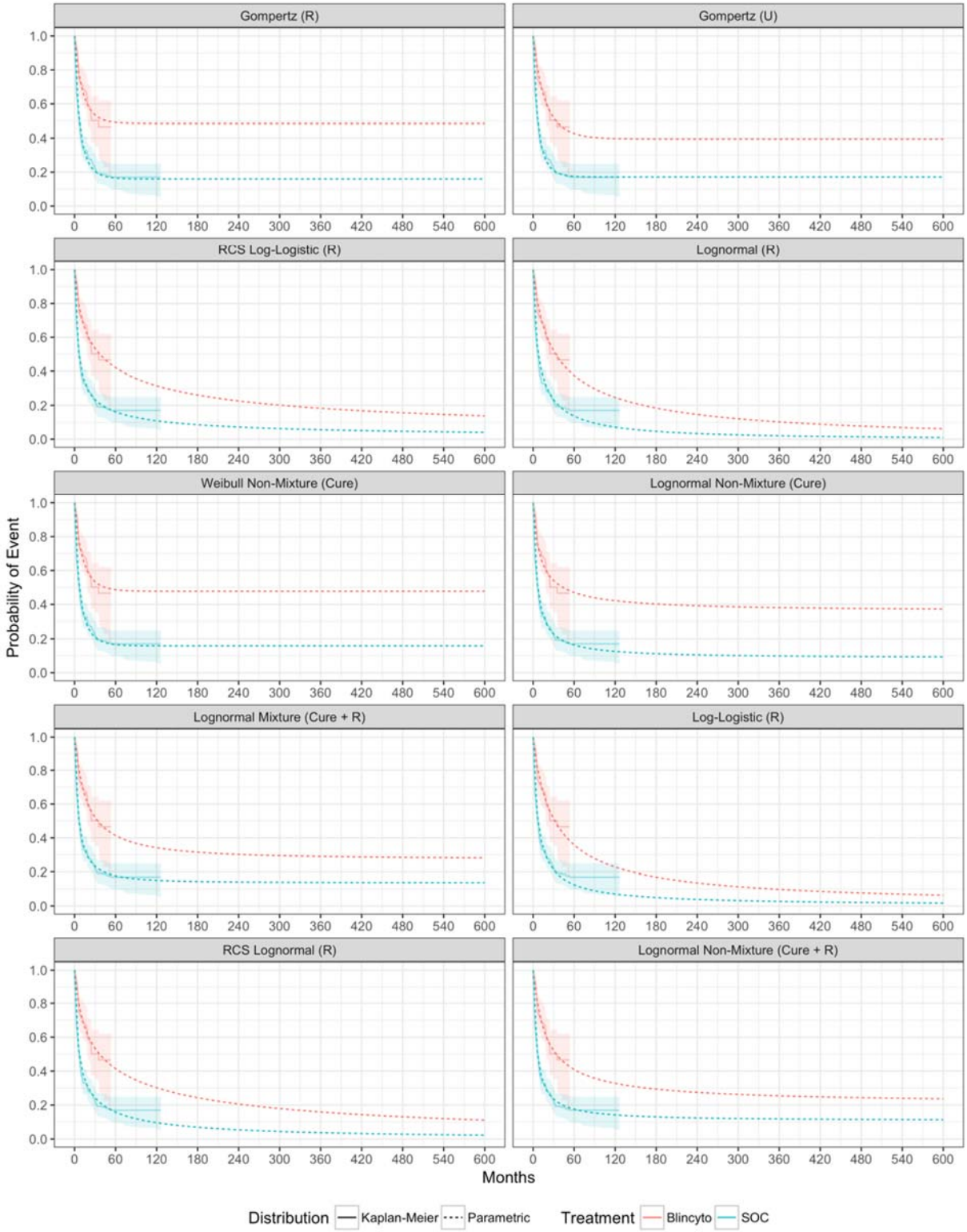
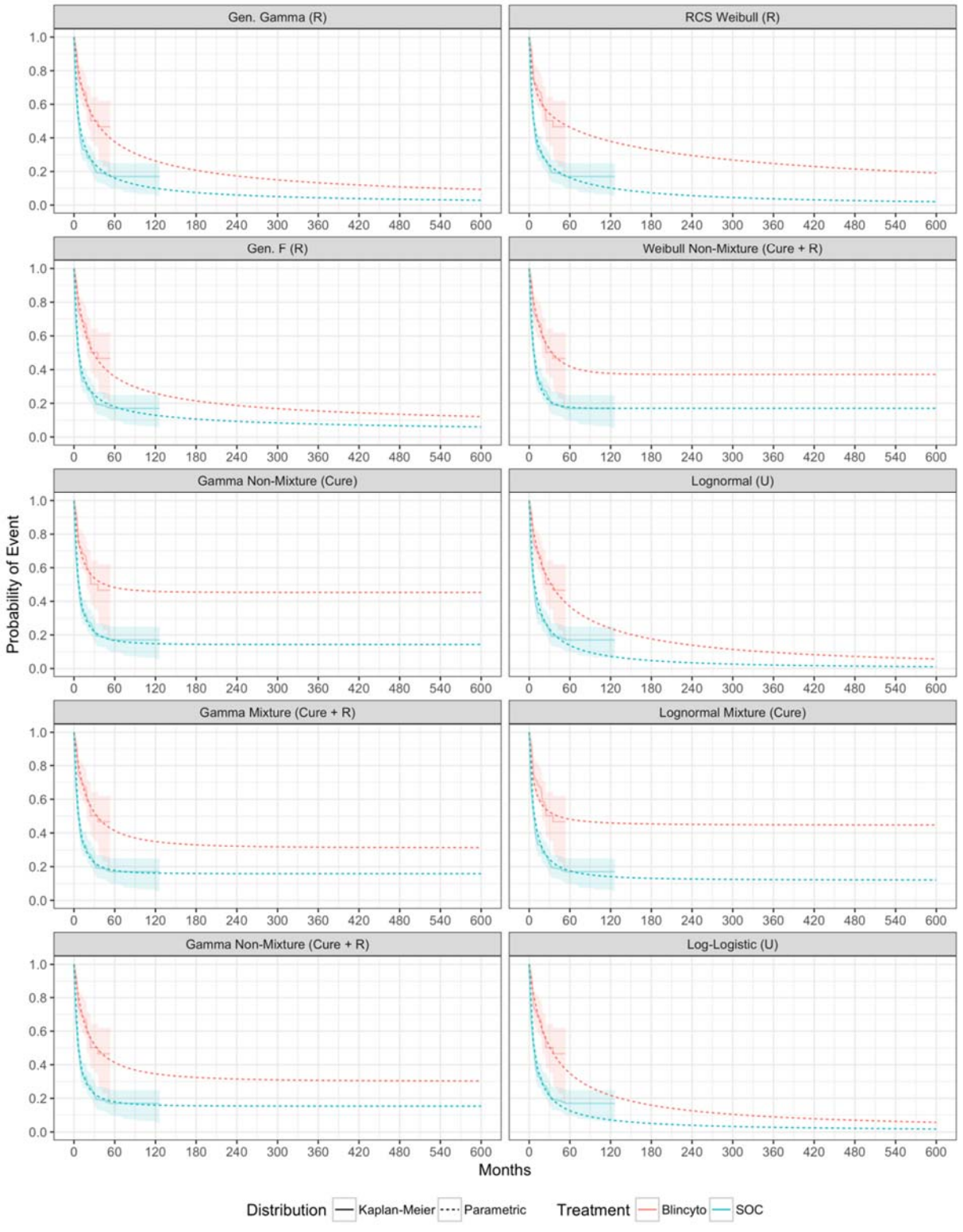
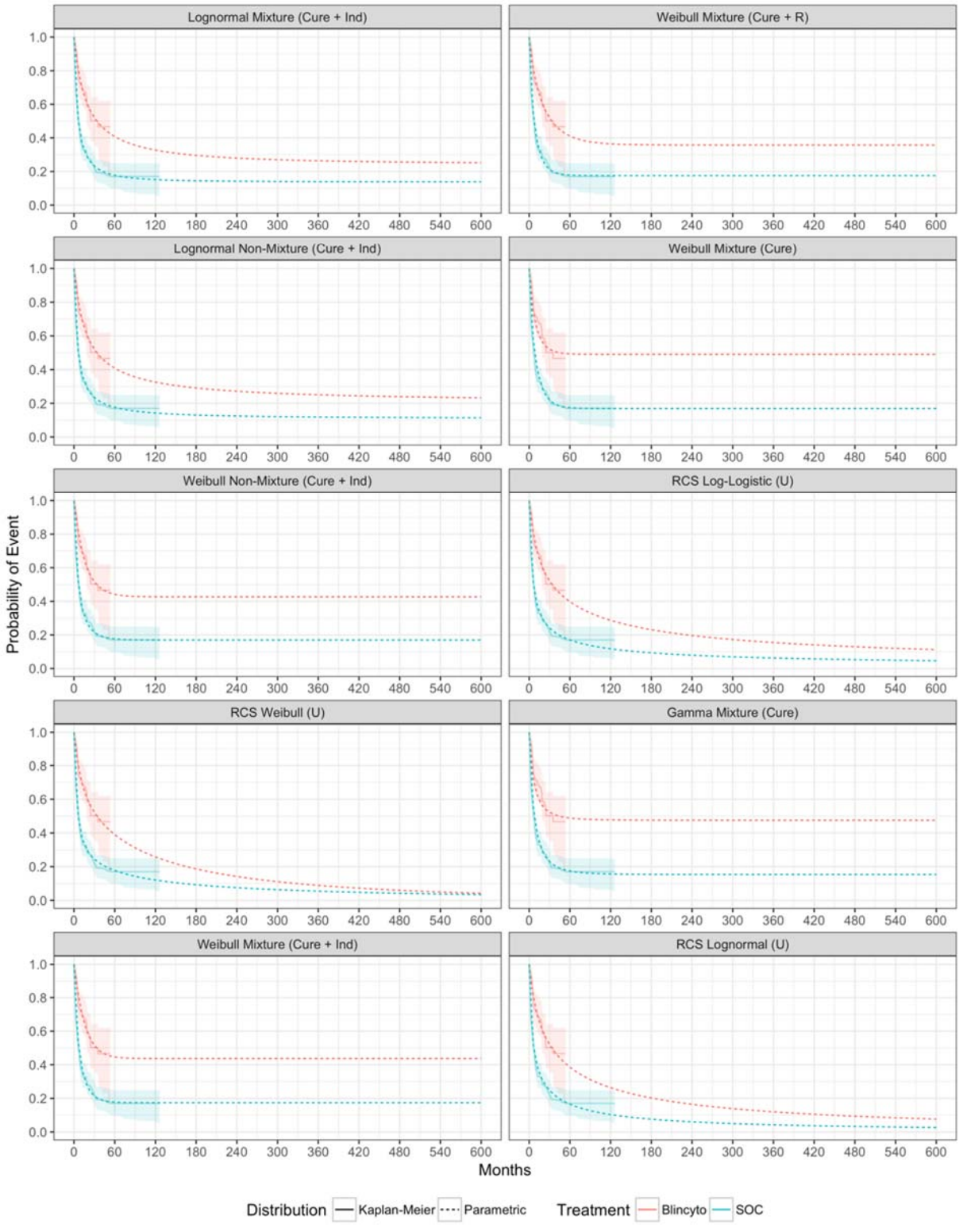

Gen. Gamma (U)
Gamma Mixture (Cure + Ind)
Gamma Non-Mixture (Cure + Ind)
Gen. F (U)
Weibull (R)
Weibull (U)
Pw. Exp.
Exponential
Probability of Event
Months
0.0
0.2
0.4
0.6
0.8
1.0
0
24
48
72
96
120
144
Distribution
Kaplan-Meier
Parametric
Treatment
Blincyto
SOC

Figure 62. Survival probabilities to 50 years for all parametric survival distributions fit to RFS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATT weights




Gen. Gamma (R)
RCS Weibull (R)
Gen. F (R)
Weibull Non-Mixture (Cure + R)
Gamma Non-Mixture (Cure)
Lognormal (U)
Gamma Mixture (Cure + R)
Lognormal Mixture (Cure)
Gamma Non-Mixture (Cure + R)
Log-Logistic (U)
Probability of Event
0.0
0.2
0.4
0.6
0.8
1.0
0
60
120
180
240
300
360
420
480
540
600
Months
Distribution
Kaplan-Meier
Parametric
Treatment
Blincyto
SOC


Lognormal Mixture (Cure + Ind)
Weibull Mixture (Cure + R)
Lognormal Non-Mixture (Cure + Ind)
Weibull Mixture (Cure)
Weibull Non-Mixture (Cure + Ind)
RCS Log-Logistic (U)
RCS Weibull (U)
Gamma Mixture (Cure)
Weibull Mixture (Cure + Ind)
RCS Lognormal (U)
Probability of Event
0.0
0.2
0.4
0.6
0.8
1.0
0
60
120
180
240
300
360
420
480
540
600
Months
Distribution
Kaplan-Meier
Parametric
Treatment
Blincyto
SOC

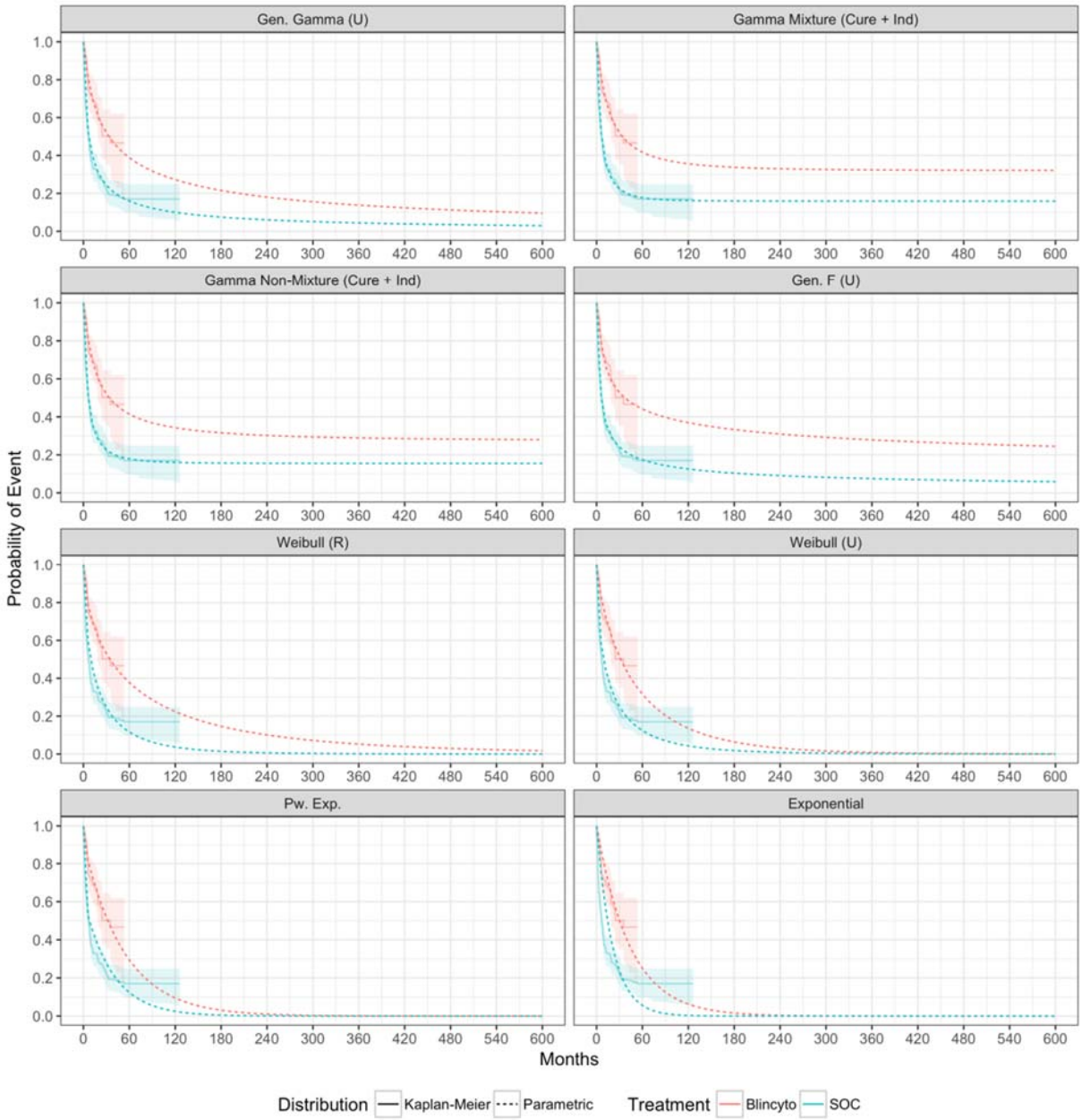


Table 90. Estimated cure fractions for all parametric survival distributions fit to RFS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATT weights

Distribution	Type	Blinatumomab	SOC
Gompertz	Restricted	48.5%	16.0%
Gompertz	Unrestricted	39.5%	17.2%
RCS Log-Logistic	Restricted	0.0%	0.0%
Lognormal	Restricted	0.0%	0.0%

Weibull Non-Mixture	Cure	47.8%	15.8%
Lognormal Non-Mixture	Cure	36.3%	8.8%
Lognormal Mixture	Cure + Restricted	27.6%	13.7%
Log-Logistic	Restricted	0.0%	0.0%
RCS Lognormal	Restricted	0.0%	0.0%
Lognormal Non-Mixture	Cure + Restricted	21.4%	10.9%
Gen. Gamma	Restricted	0.0%	0.0%
RCS Weibull	Restricted	0.0%	0.0%
Gen. F	Restricted	0.0%	0.0%
Weibull Non-Mixture	Cure + Restricted	37.1%	17.1%
Gamma Non-Mixture	Cure	45.3%	14.3%
Lognormal	Unrestricted	0.0%	0.0%
Gamma Mixture	Cure	31.2%	15.8%
Lognormal Mixture	Cure	44.7%	12.0%
Gamma Non-Mixture	Cure + Restricted	30.2%	15.4%
Log-Logistic	Unrestricted	0.0%	0.0%
Lognormal Mixture	Cure + Unrestricted	24.3%	13.8%
Weibull Mixture	Cure + Restricted	35.7%	17.5%
Lognormal Non-Mixture	Cure + Unrestricted	20.8%	11.0%
Weibull Mixture	Cure	49.1%	16.9%
Weibull Non-Mixture	Cure + Unrestricted	42.7%	17.0%
RCS Log-Logistic	Unrestricted	0.0%	0.0%
RCS Weibull	Unrestricted	0.0%	0.0%
Gamma Mixture	Cure	47.5%	15.3%
Weibull Mixture	Cure + Unrestricted	43.7%	17.4%
RCS Lognormal	Unrestricted	0.0%	0.0%
Gen. Gamma	Unrestricted	0.0%	0.0%
Gamma Mixture	Cure + Unrestricted	32.0%	15.9%
Gamma Non-Mixture	Cure + Unrestricted	27.3%	15.5%
Gen. F	Unrestricted	0.0%	0.0%
Weibull	Restricted	0.0%	0.0%
Weibull	Unrestricted	0.0%	0.0%
Pw. Exp.	Unrestricted	0.0%	0.0%
Exponential	Restricted	0.0%	0.0%

1.1.2 Overall Survival

Figure 63. Kaplan Meier estimates of OS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATT weights

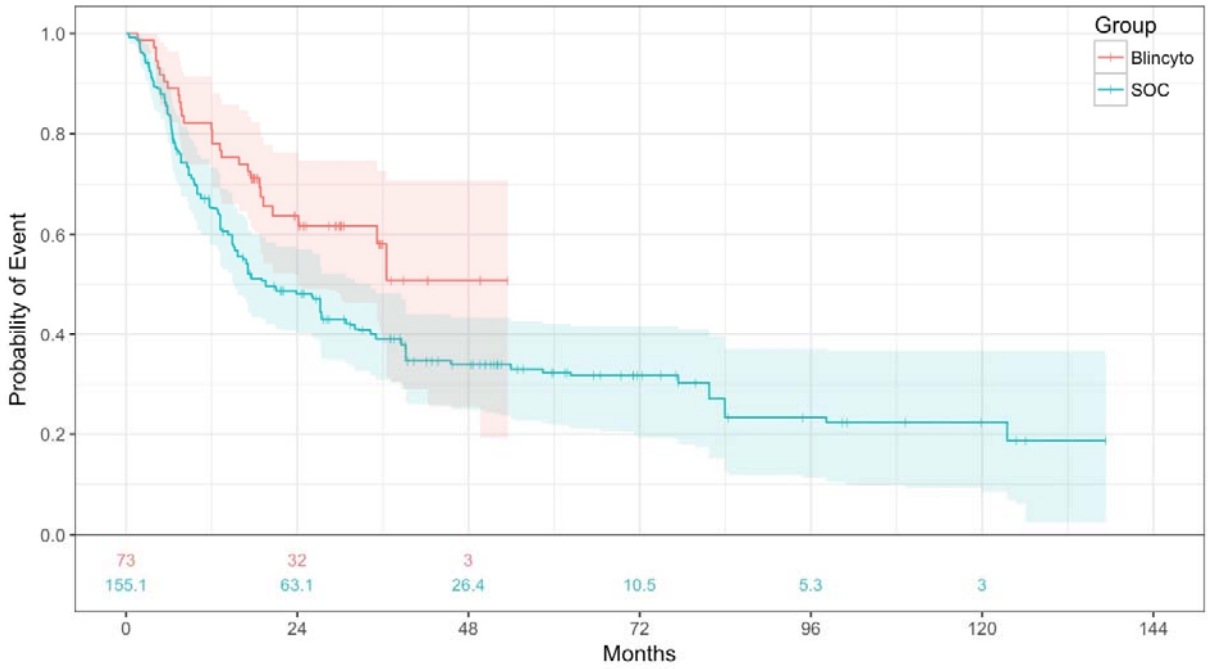


Figure 64. Fit statistics all parametric distributions fit to OS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATT weights

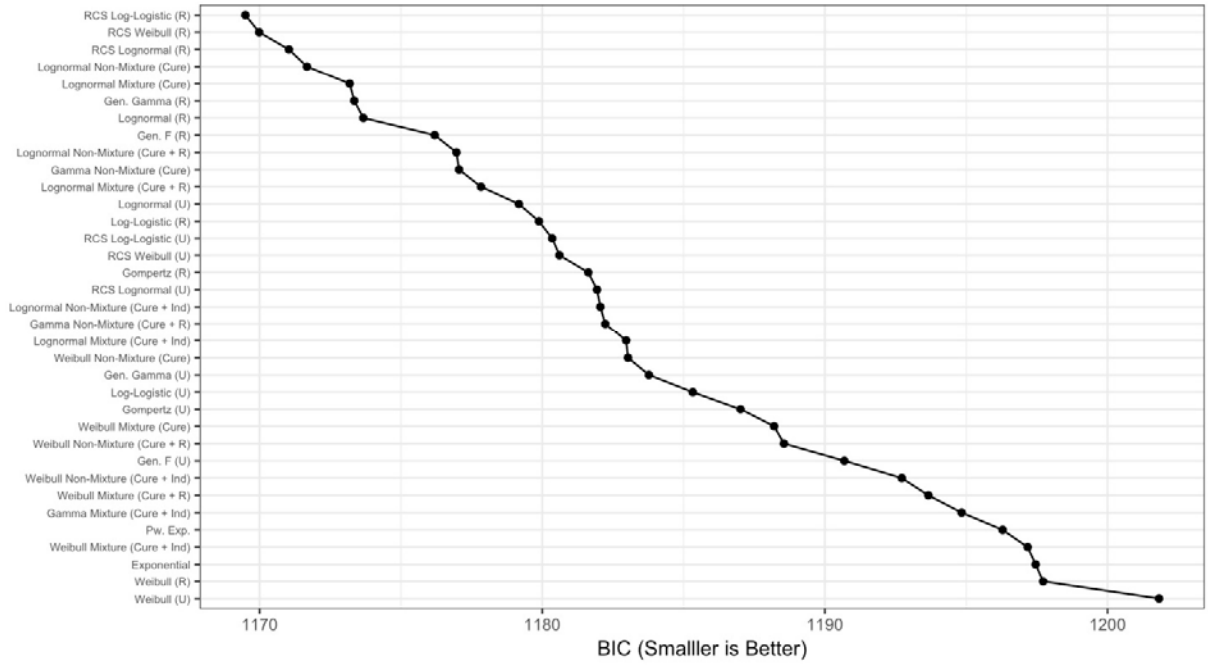


Figure 65. Treatment effect counterfactual plots for OS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATT weights

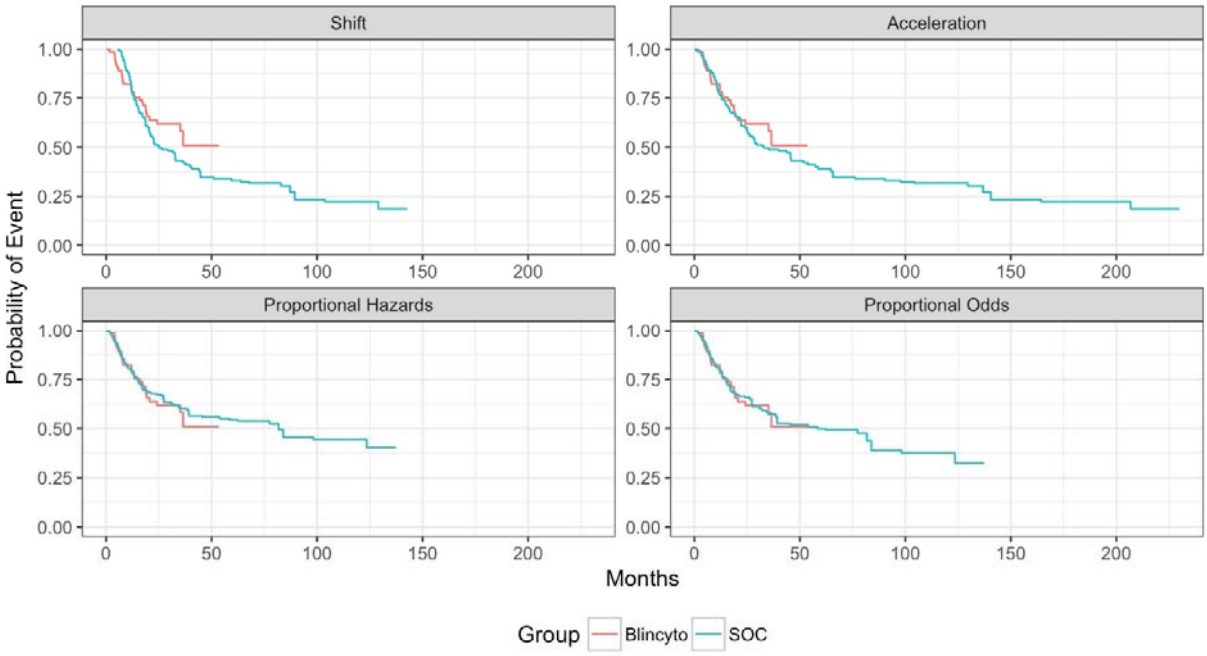


Figure 66. Schoenfeld residuals for OS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATT weights

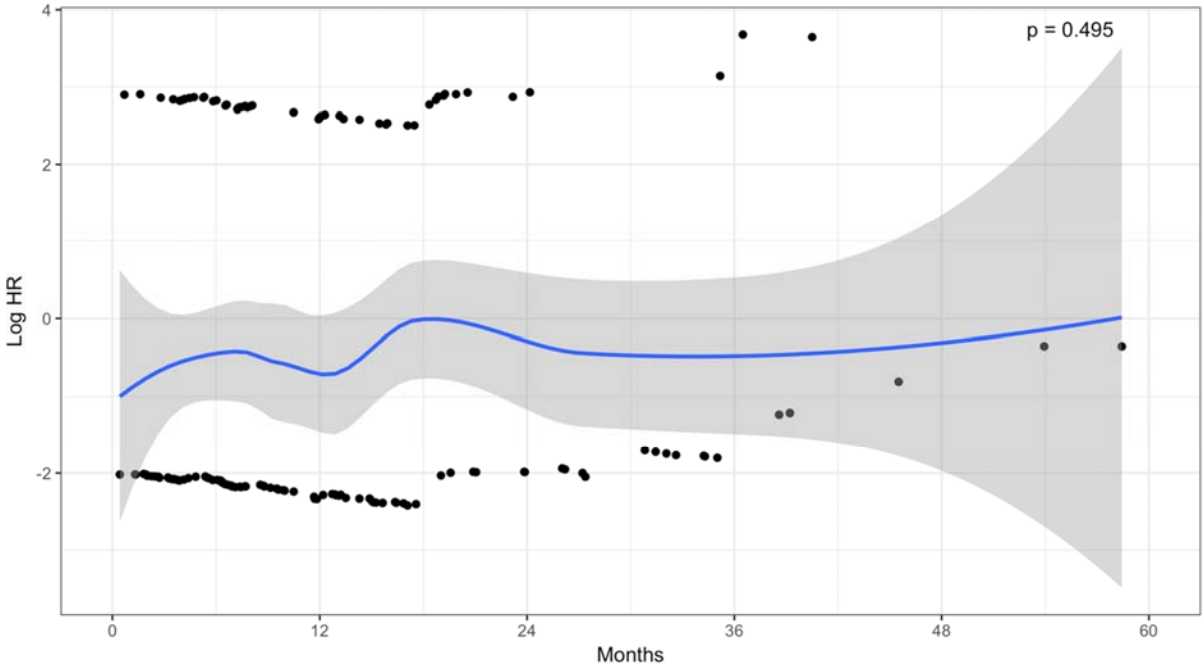
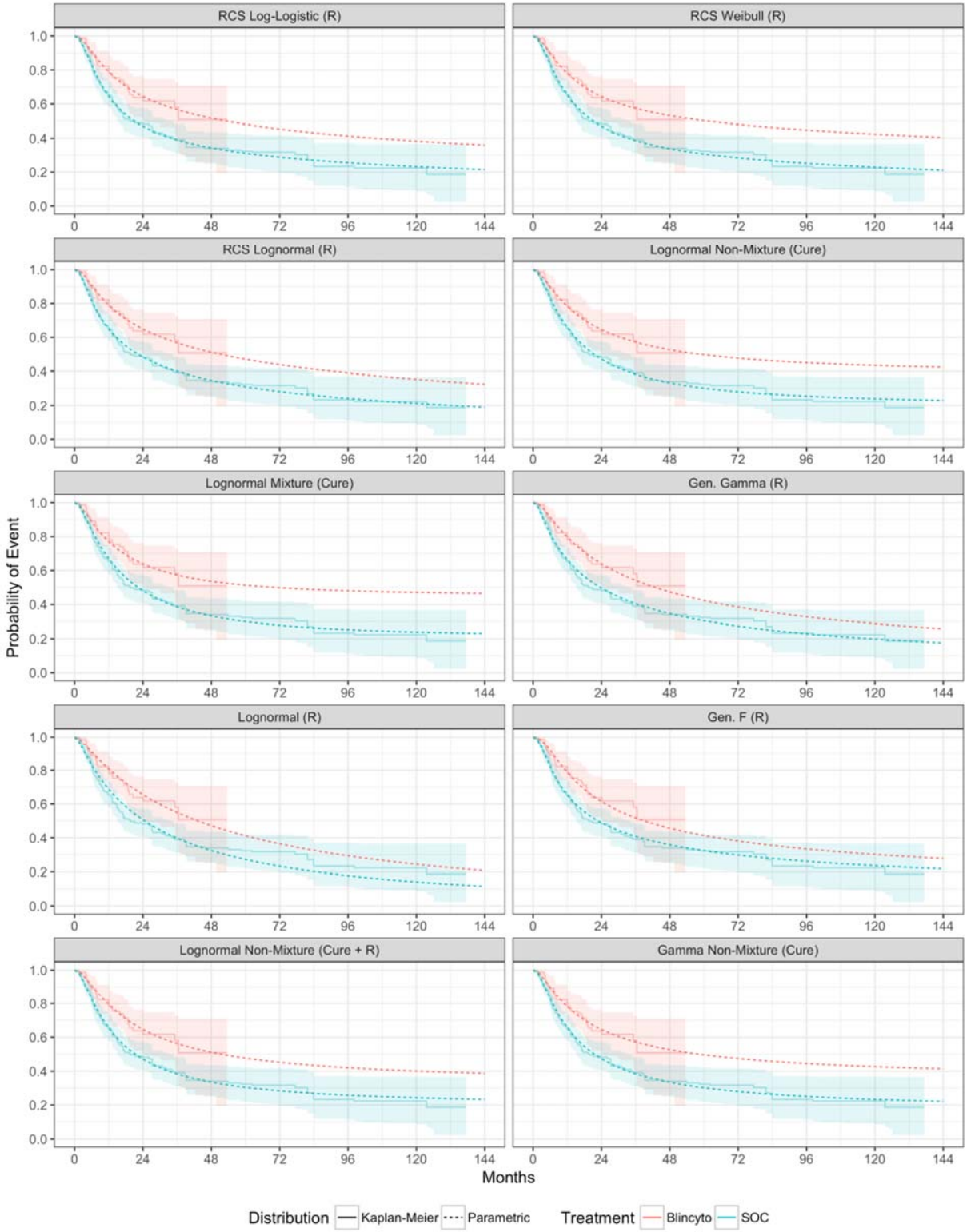
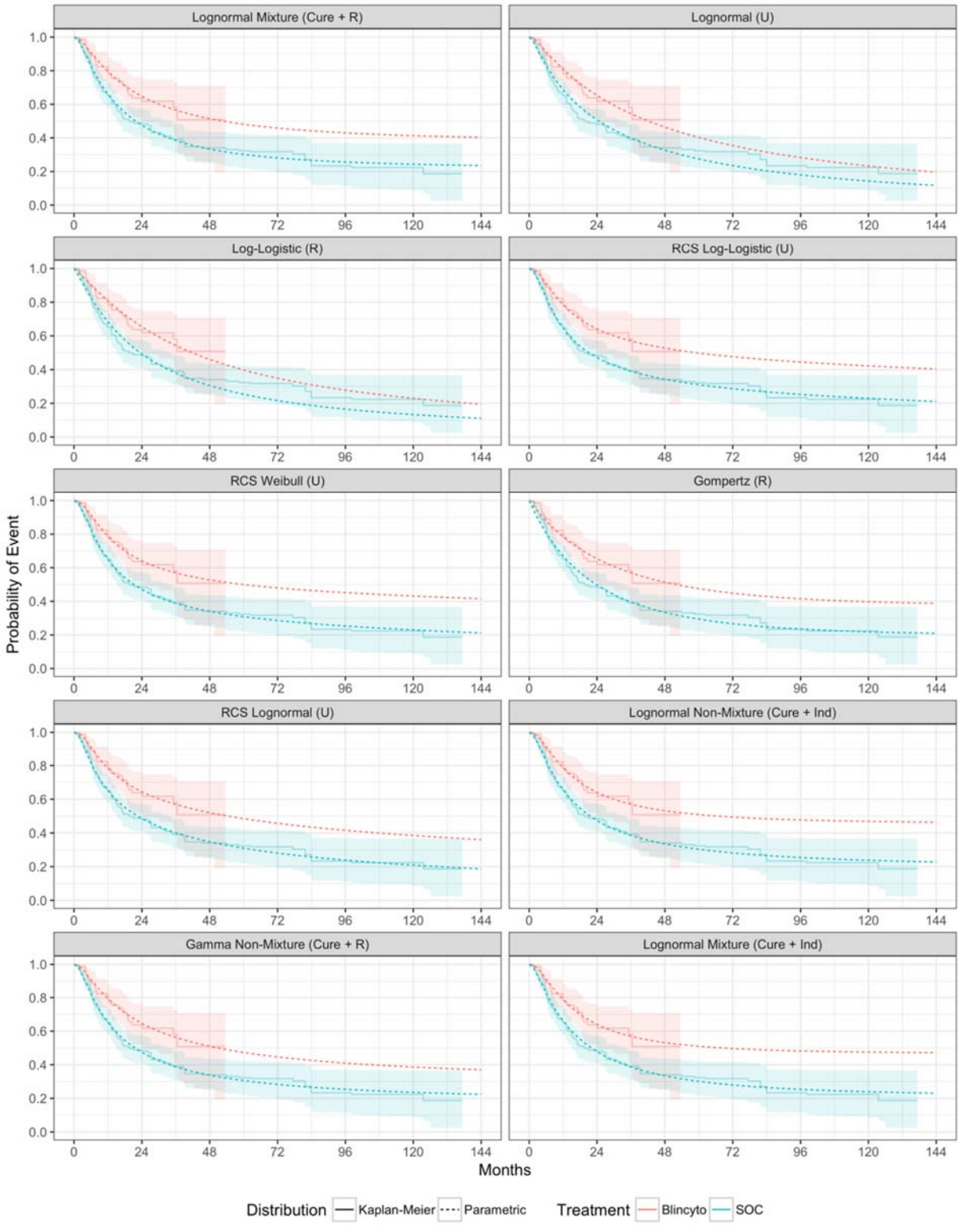
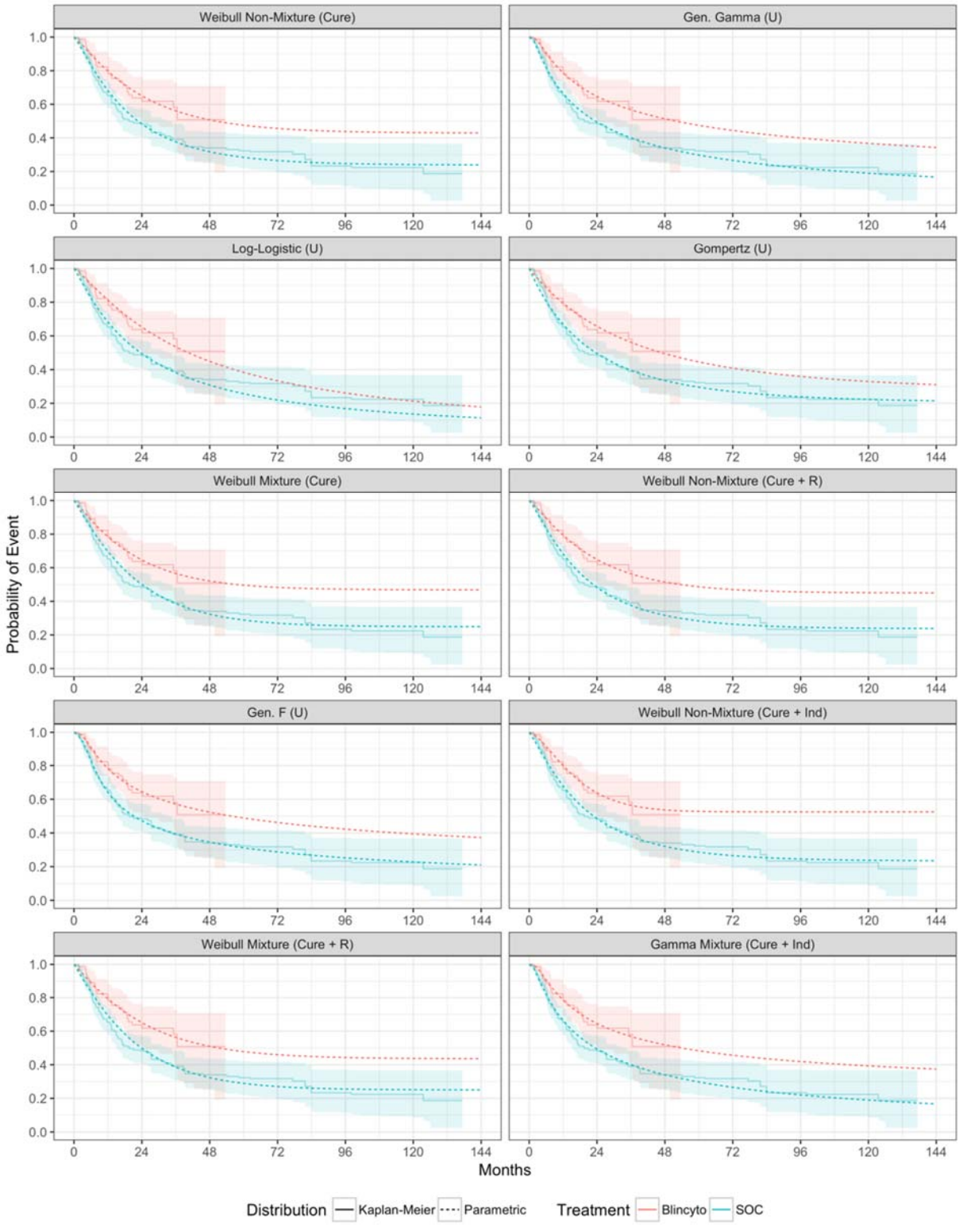


Figure 67. Survival probabilities to 12 years for all parametric distributions fit to OS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATT weights




Lognormal Mixture (Cure + R)
Lognormal (U)
Log-Logistic (R)
RCS Log-Logistic (U)
RCS Weibull (U)
Gompertz (R)
RCS Lognormal (U)
Lognormal Non-Mixture (Cure + Ind)
Gamma Non-Mixture (Cure + R)
Lognormal Mixture (Cure + Ind)
Probability of Event
0.0
0.2
0.4
0.6
0.8
1.0
0
24
48
72
96
120
144
Months
Distribution
Kaplan-Meier
Parametric
Treatment
Blincyto
SOC


Weibull Non-Mixture (Cure)
Gen. Gamma (U)
Log-Logistic (U)
Gompertz (U)
Weibull Mixture (Cure)
Weibull Non-Mixture (Cure + R)
Gen. F (U)
Weibull Non-Mixture (Cure + Ind)
Weibull Mixture (Cure + R)
Gamma Mixture (Cure + Ind)
Probability of Event
0.0
0.2
0.4
0.6
0.8
1.0
0
24
48
72
96
120
144
Months
Distribution
Kaplan-Meier
Parametric
Treatment
Blincyto
SOC

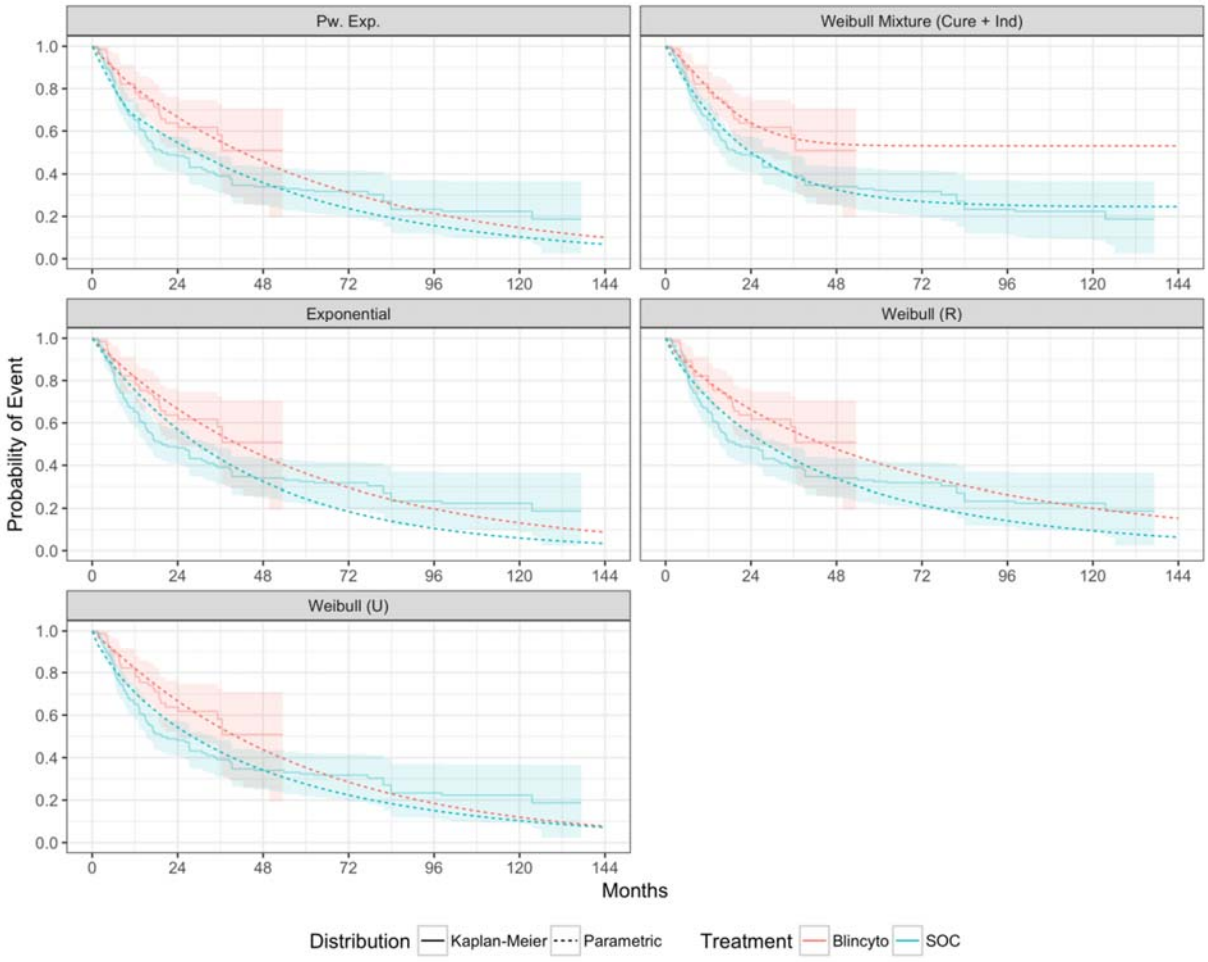
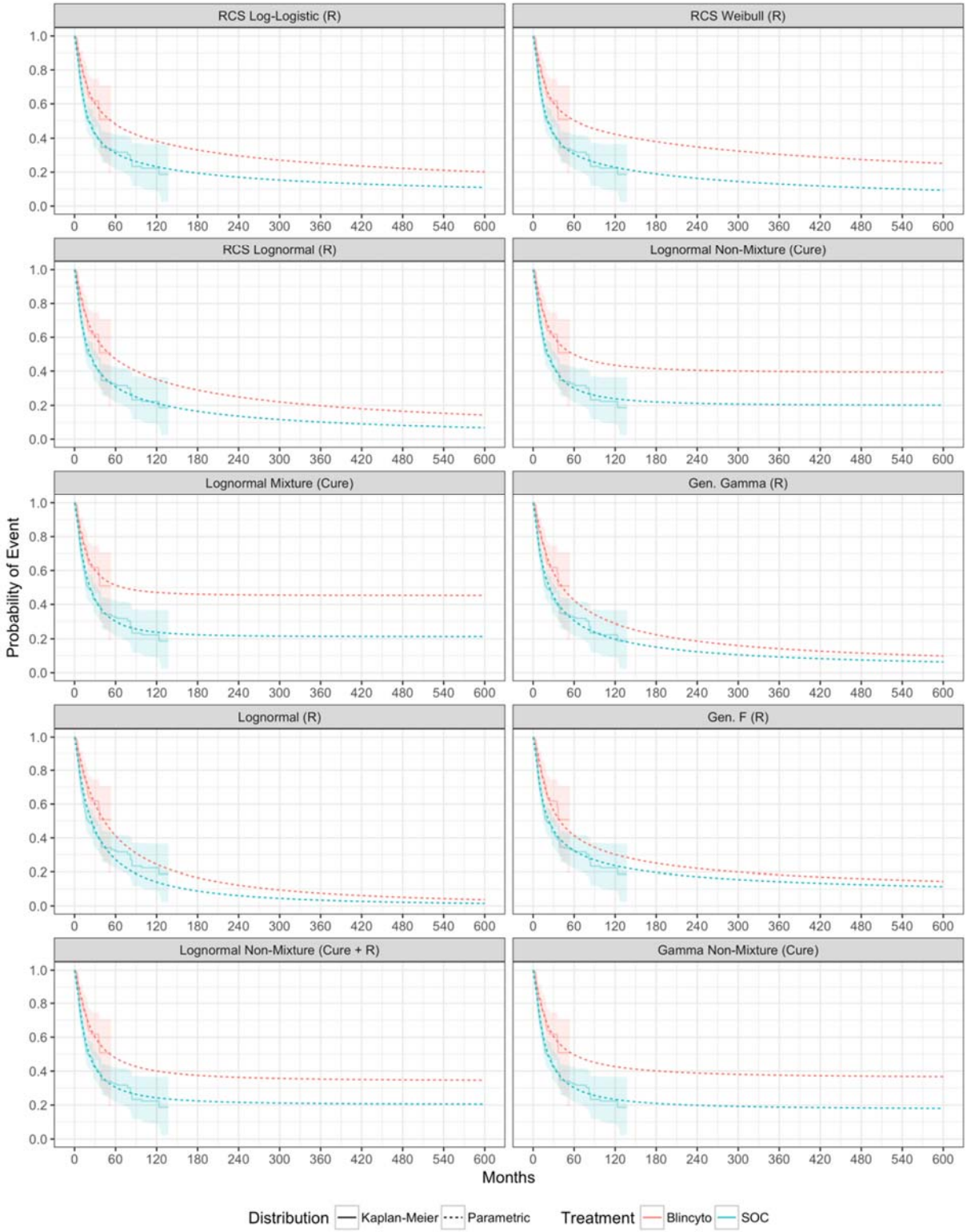
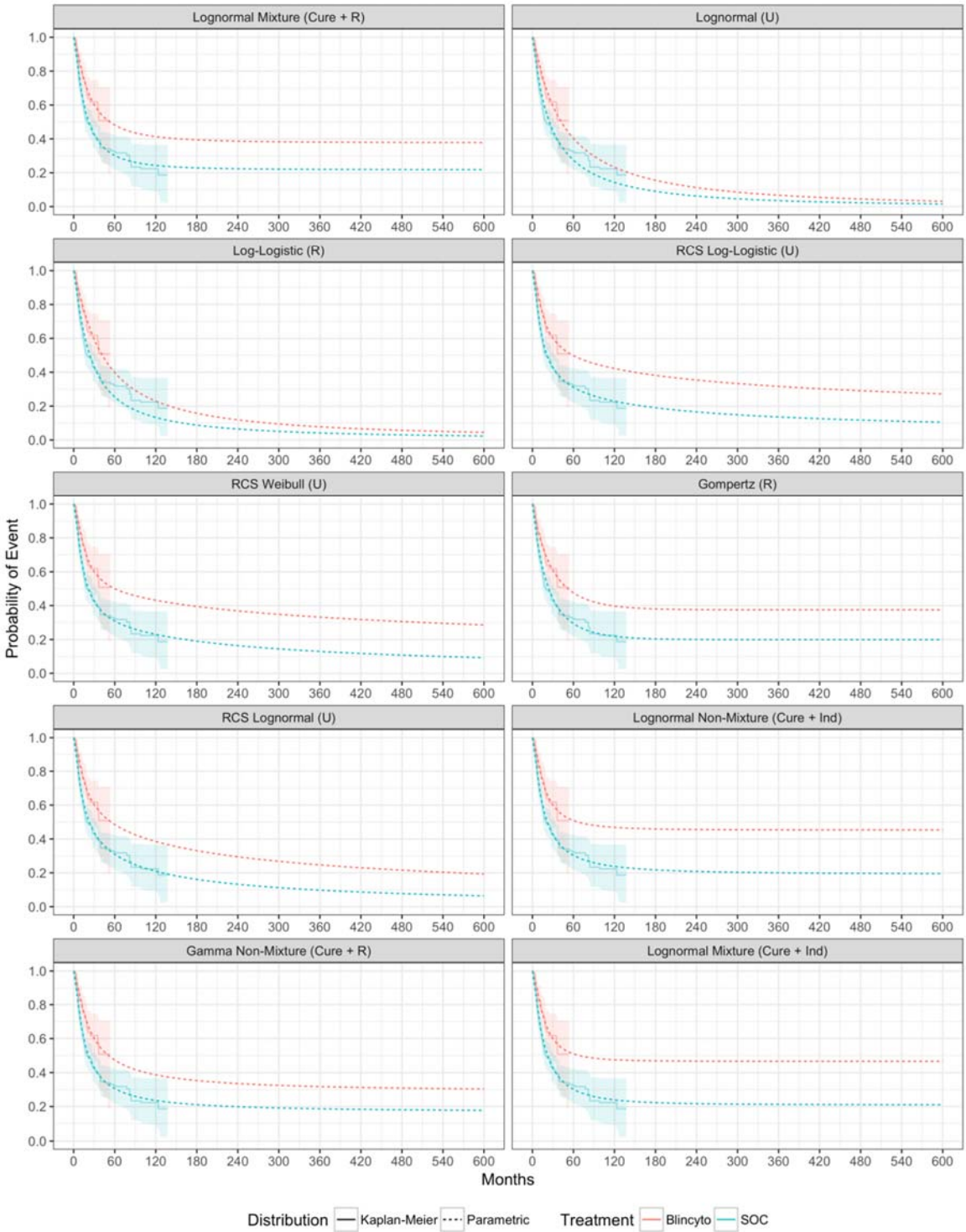
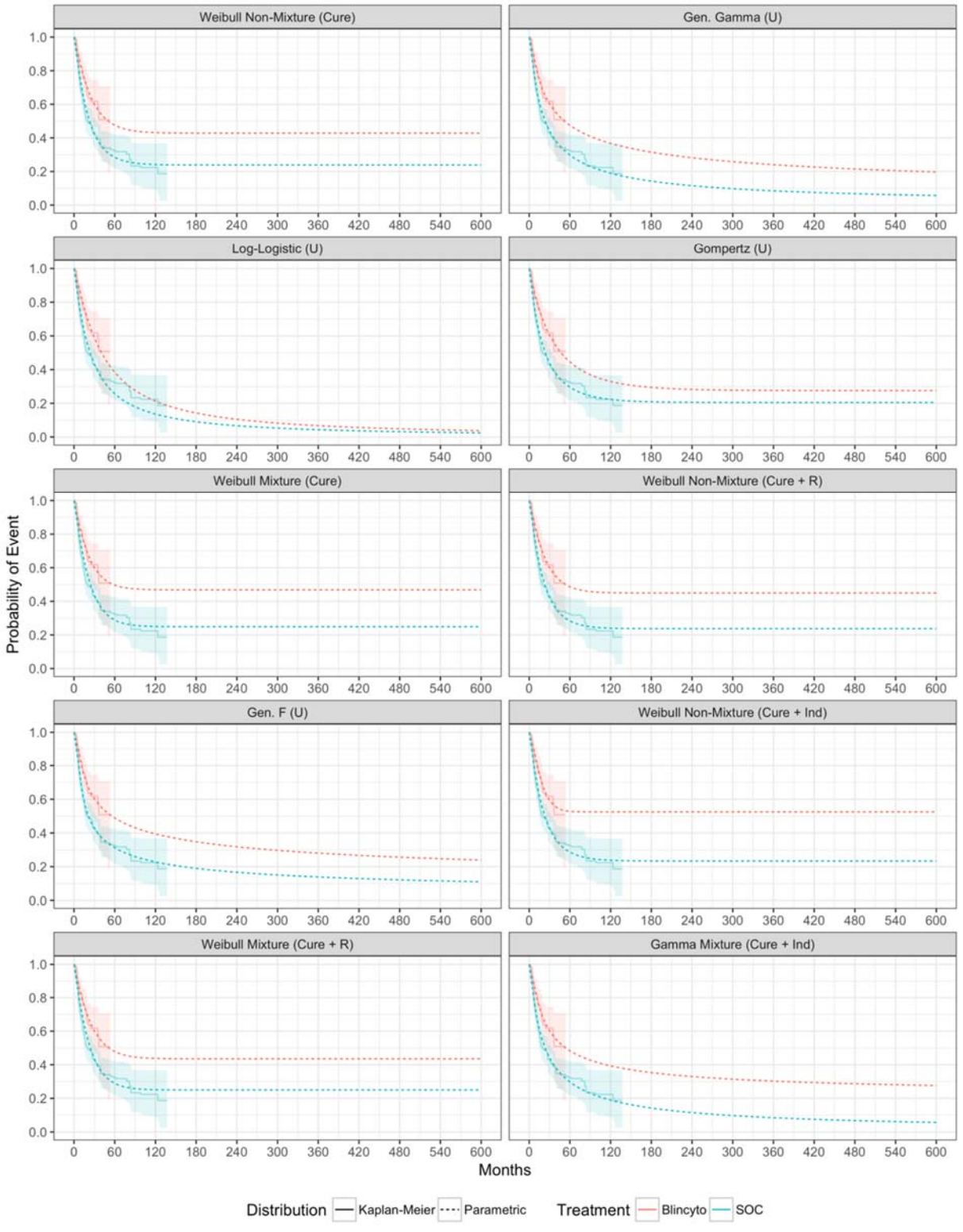

Pw. Exp.
Weibull Mixture (Cure + Ind)
Exponential
Weibull (R)
Weibull (U)
Probability of Event
1.0
0.8
0.6
0.4
0.2
0.0
0
24
48
72
96
120
144
Months
Distribution
Kaplan-Meier
Parametric
Treatment
Blincyto
SOC

Figure 68. Survival probabilities to 50 years for all parametric distributions fit to OS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATT weights




Lognormal Mixture (Cure + R)
Lognormal (U)
Log-Logistic (R)
RCS Log-Logistic (U)
RCS Weibull (U)
Gompertz (R)
RCS Lognormal (U)
Lognormal Non-Mixture (Cure + Ind)
Gamma Non-Mixture (Cure + R)
Lognormal Mixture (Cure + Ind)
Probability of Event
0.0
0.2
0.4
0.6
0.8
1.0
0
60
120
180
240
300
360
420
480
540
600
Months
Distribution
Kaplan-Meier
Parametric
Treatment
Blincyto
SOC


Weibull Non-Mixture (Cure)
Gen. Gamma (U)
Log-Logistic (U)
Gompertz (U)
Weibull Mixture (Cure)
Weibull Non-Mixture (Cure + R)
Gen. F (U)
Weibull Non-Mixture (Cure + Ind)
Weibull Mixture (Cure + R)
Gamma Mixture (Cure + Ind)
Probability of Event
0.0
0.2
0.4
0.6
0.8
1.0
0
60
120
180
240
300
360
420
480
540
600
Months
Distribution
Kaplan-Meier
Parametric
Treatment
Blincyto
SOC

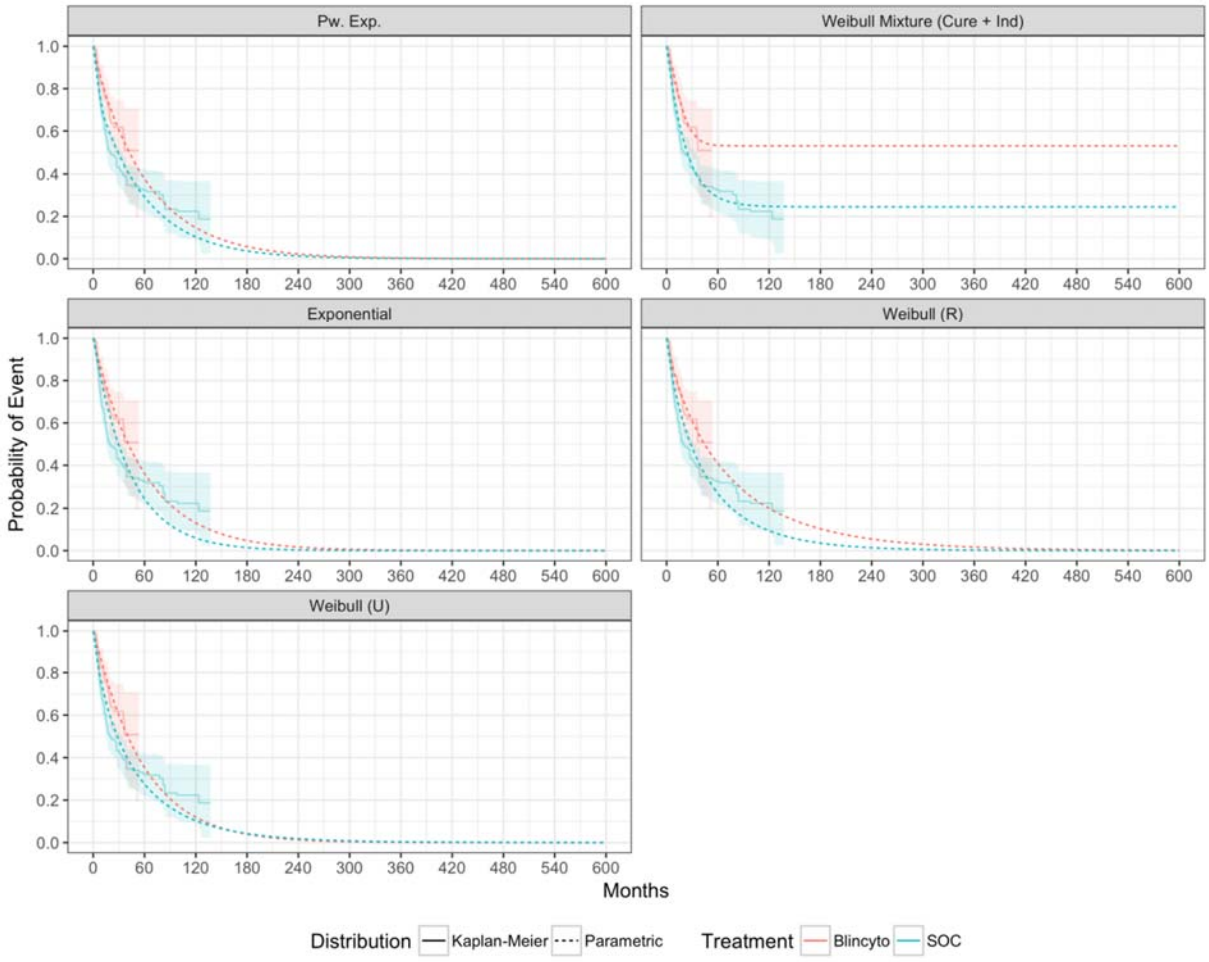


Table 91. Estimated cure fractions for all parametric distributions fit to OS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATT weights

Distribution	Type	Blinatumomab	SOC
RCS Log-Logistic	Restricted	0.0%	0.0%
RCS Weibull	Restricted	0.0%	0.0%
RCS Lognormal	Restricted	0.0%	0.0%
Lognormal Non-Mixture	Cure	39.1%	19.9%
Lognormal Mixture	Cure	45.3%	21.3%
Gen. Gamma	Restricted	0.0%	0.0%
Lognormal	Restricted	0.0%	0.0%
Gen. F	Restricted	0.0%	0.0%
Lognormal Non-Mixture	Cure + Restricted	34.3%	20.4%
Gamma Non-Mixture	Cure	36.0%	17.4%
Lognormal Mixture	Cure + Restricted	37.8%	21.8%
Lognormal	Unrestricted	0.0%	0.0%
Log-Logistic	Restricted	0.0%	0.0%

RCS Log-Logistic	Unrestricted	0.1%	0.0%
RCS Weibull	Unrestricted	0.0%	0.0%
Gompertz	Restricted	37.5%	19.9%
RCS Lognormal	Unrestricted	0.0%	0.0%
Lognormal Non-Mixture	Cure + Unrestricted	45.3%	19.3%
Gamma Non-Mixture	Cure + Restricted	28.8%	16.8%
Lognormal Mixture	Cure + Unrestricted	46.6%	21.0%
Weibull Non-Mixture	Cure	42.8%	23.8%
Gen. Gamma	Unrestricted	0.1%	0.0%
Log-Logistic	Unrestricted	0.0%	0.0%
Gompertz	Unrestricted	27.5%	20.5%
Weibull Mixture	Cure	46.8%	24.9%
Weibull Non-Mixture	Cure + Restricted	45.0%	23.7%
Gen. F	Unrestricted	0.3%	0.0%
Weibull Non-Mixture	Cure + Unrestricted	52.5%	23.3%
Weibull Mixture	Cure + Restricted	43.5%	25.0%
Gamma Mixture	Cure + Unrestricted	20.8%	0.0%
Pw. Exp.	Unrestricted	0.0%	0.0%
Weibull Mixture	Cure + Unrestricted	53.0%	24.5%
Exponential	Restricted	0.0%	0.0%
Weibull	Restricted	0.0%	0.0%
Weibull	Unrestricted	0.0%	0.0%

1.2 Relapse-Free and Overall Survival Using ATE Weights

1.2.1 Relapse-Free Survival

Figure 69. Kaplan Meier estimates of RFS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATE weights

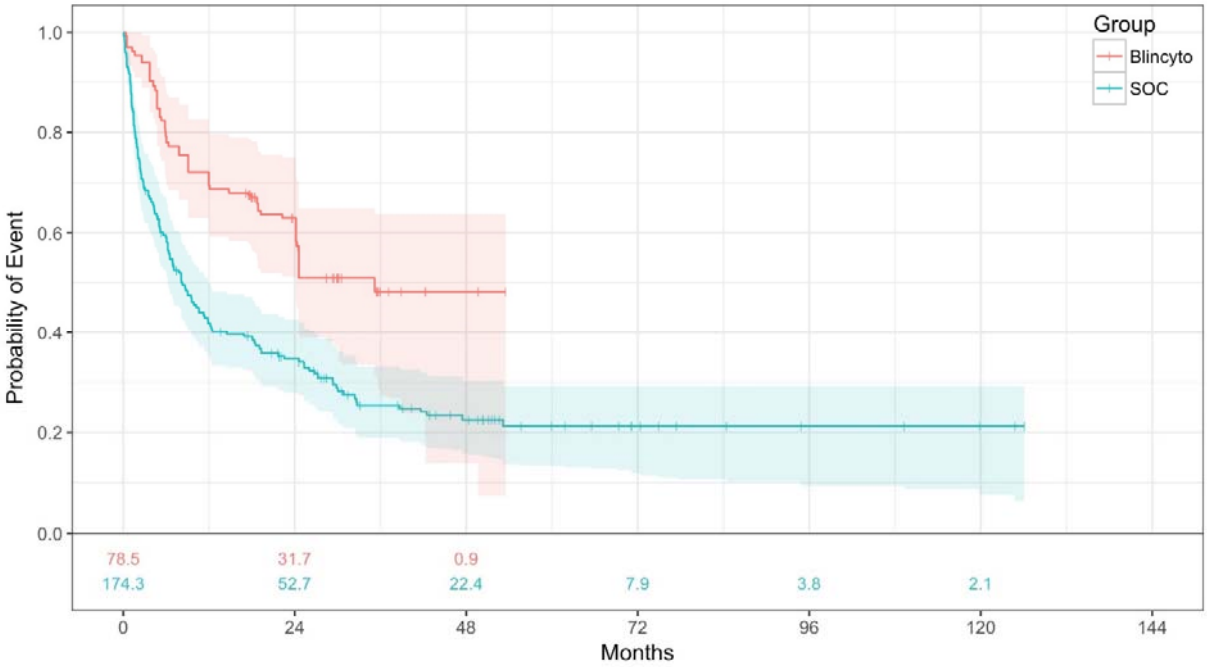


Figure 70. Fit statistics for parametric survival distributions fit to RFS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATE weights

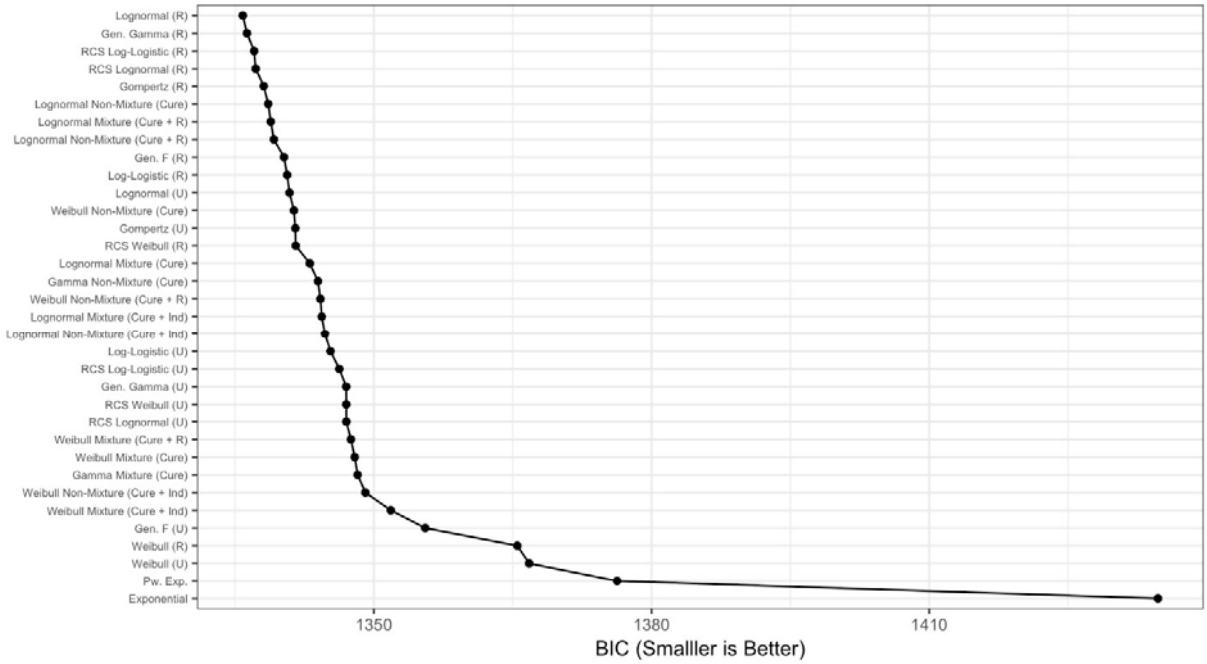


Figure 71. Treatment effect counterfactual plots for RFS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATE weights

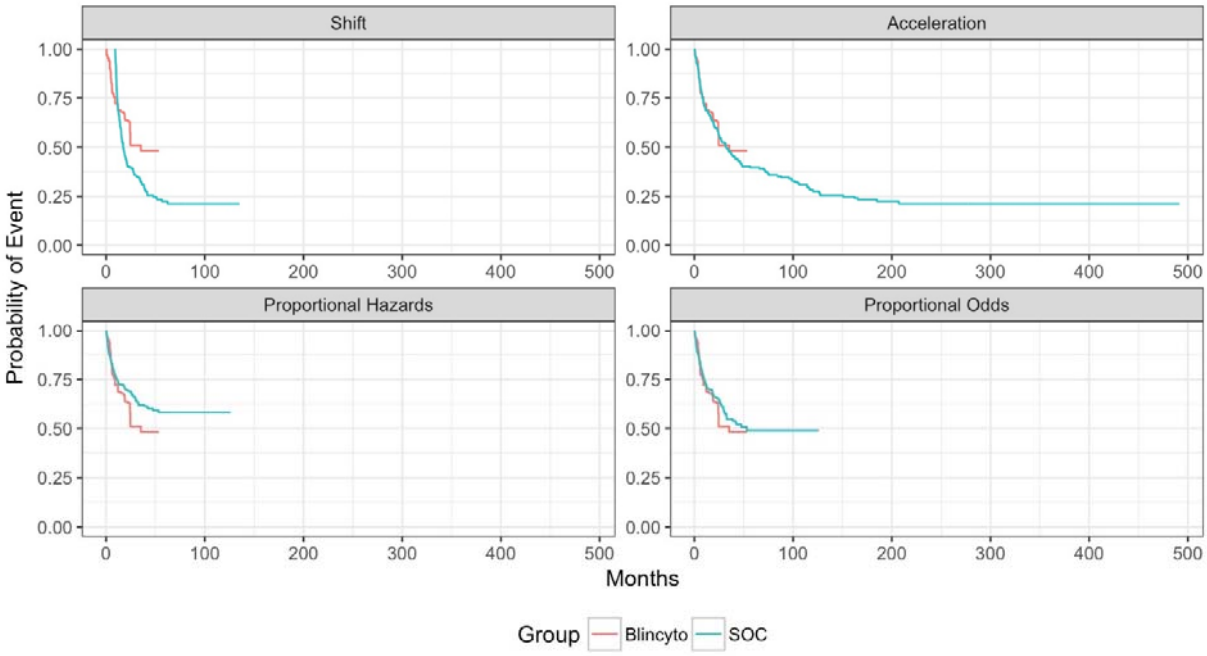


Figure 72. Schoenfeld residual plots for RFS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATE weights

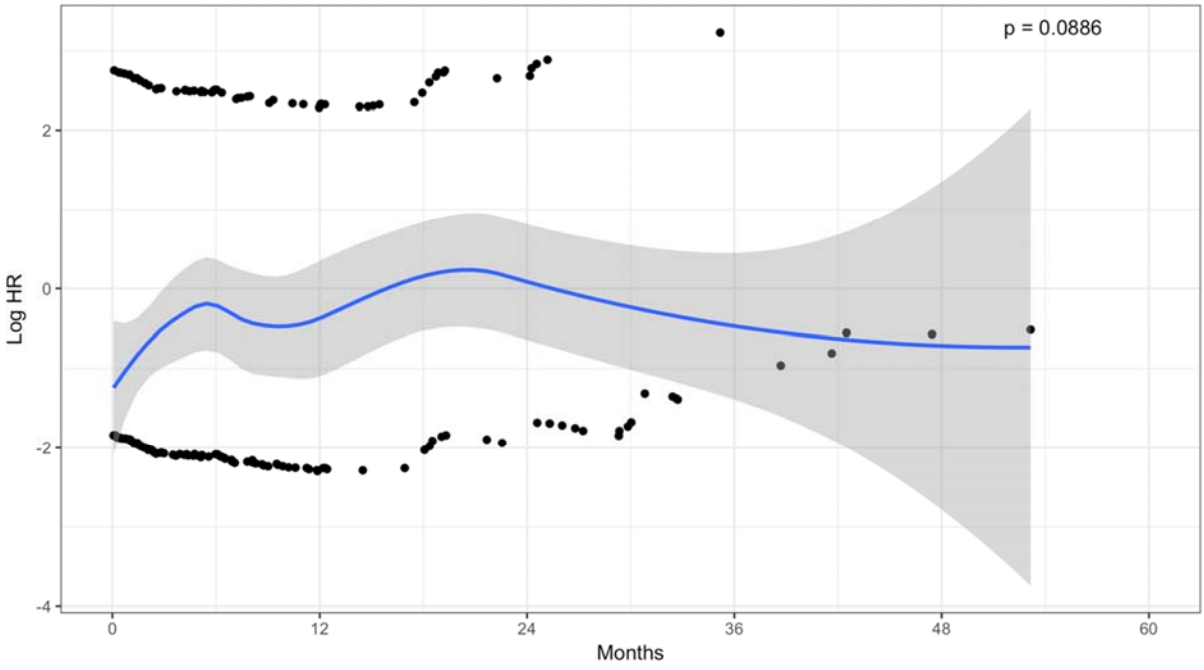
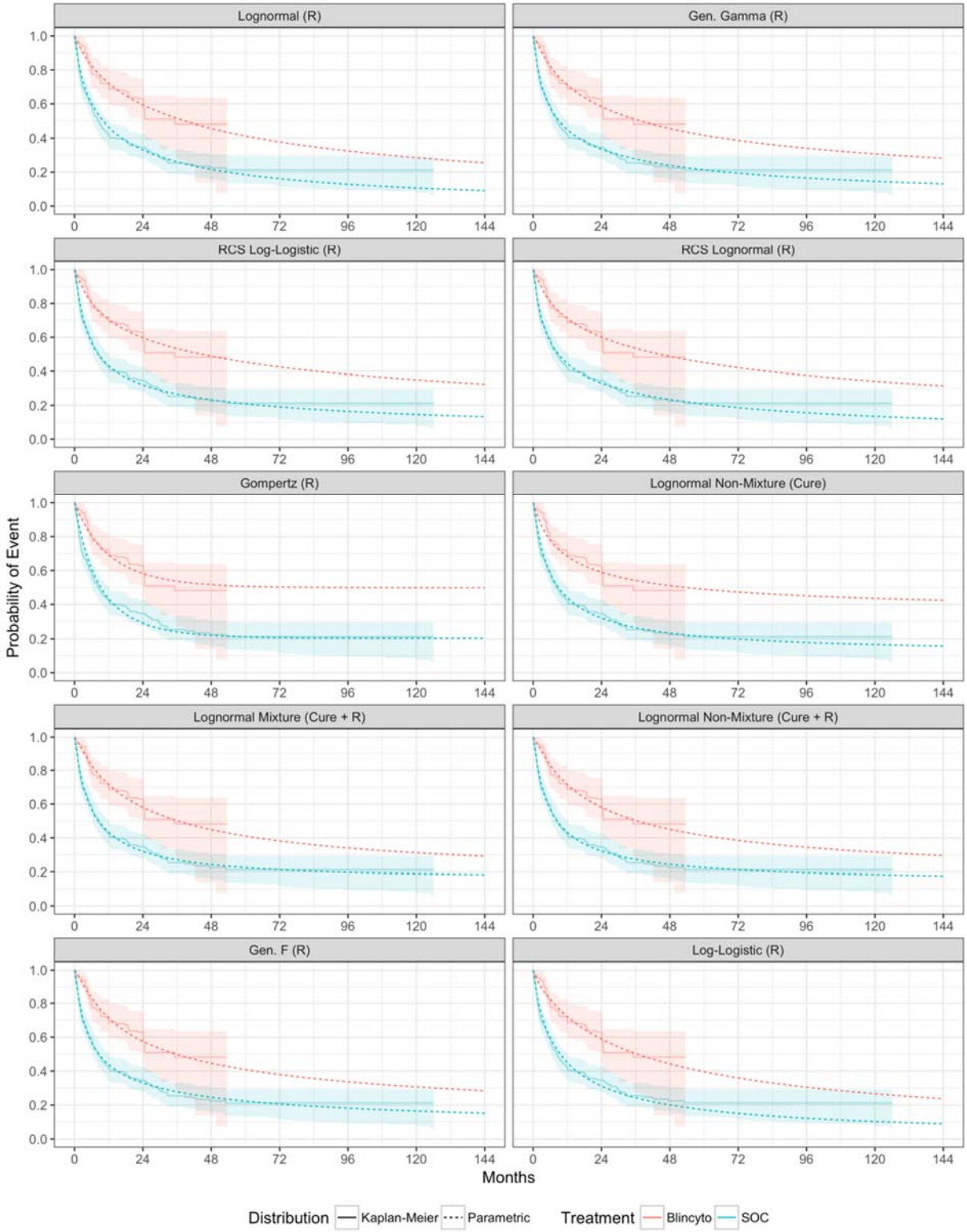
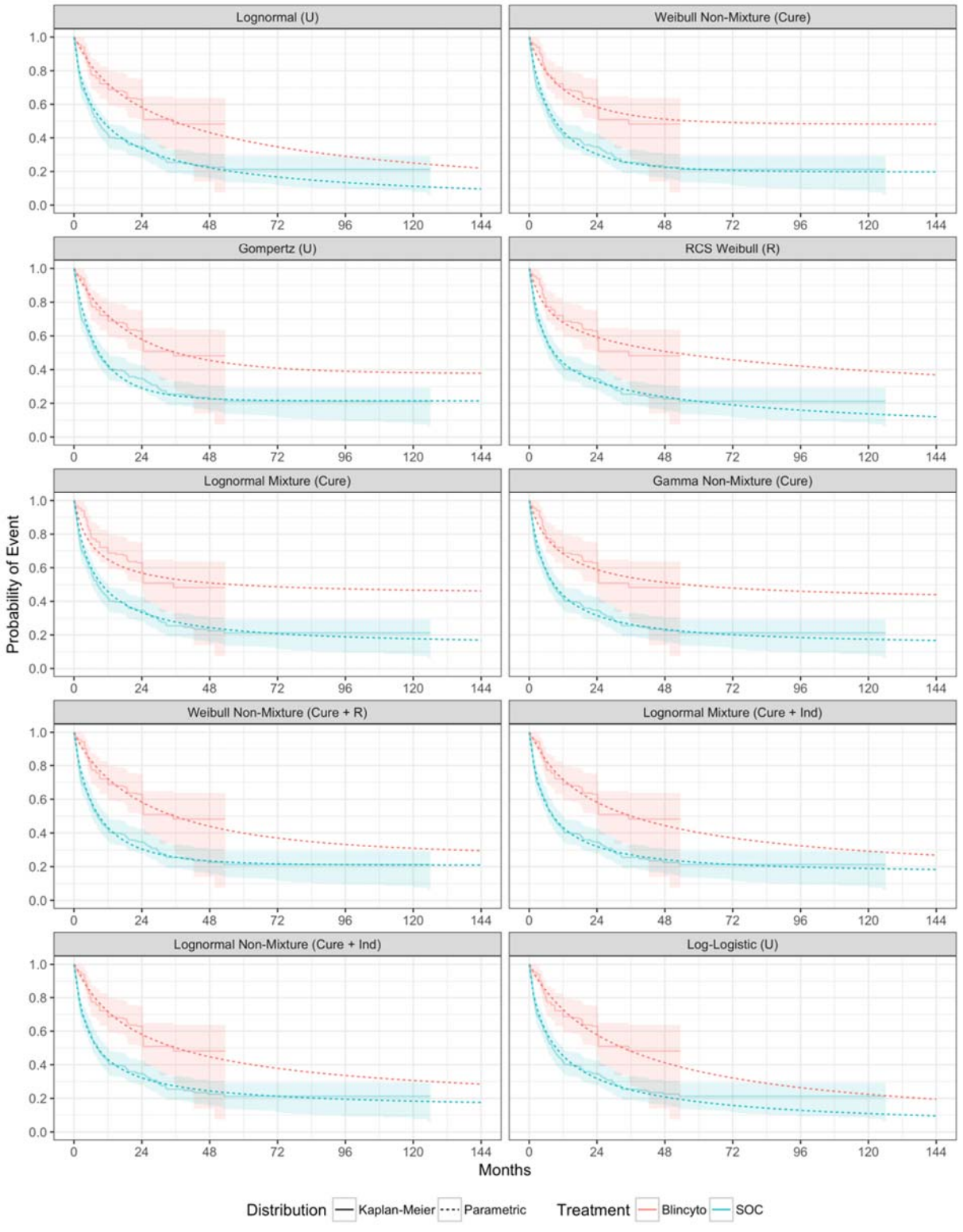
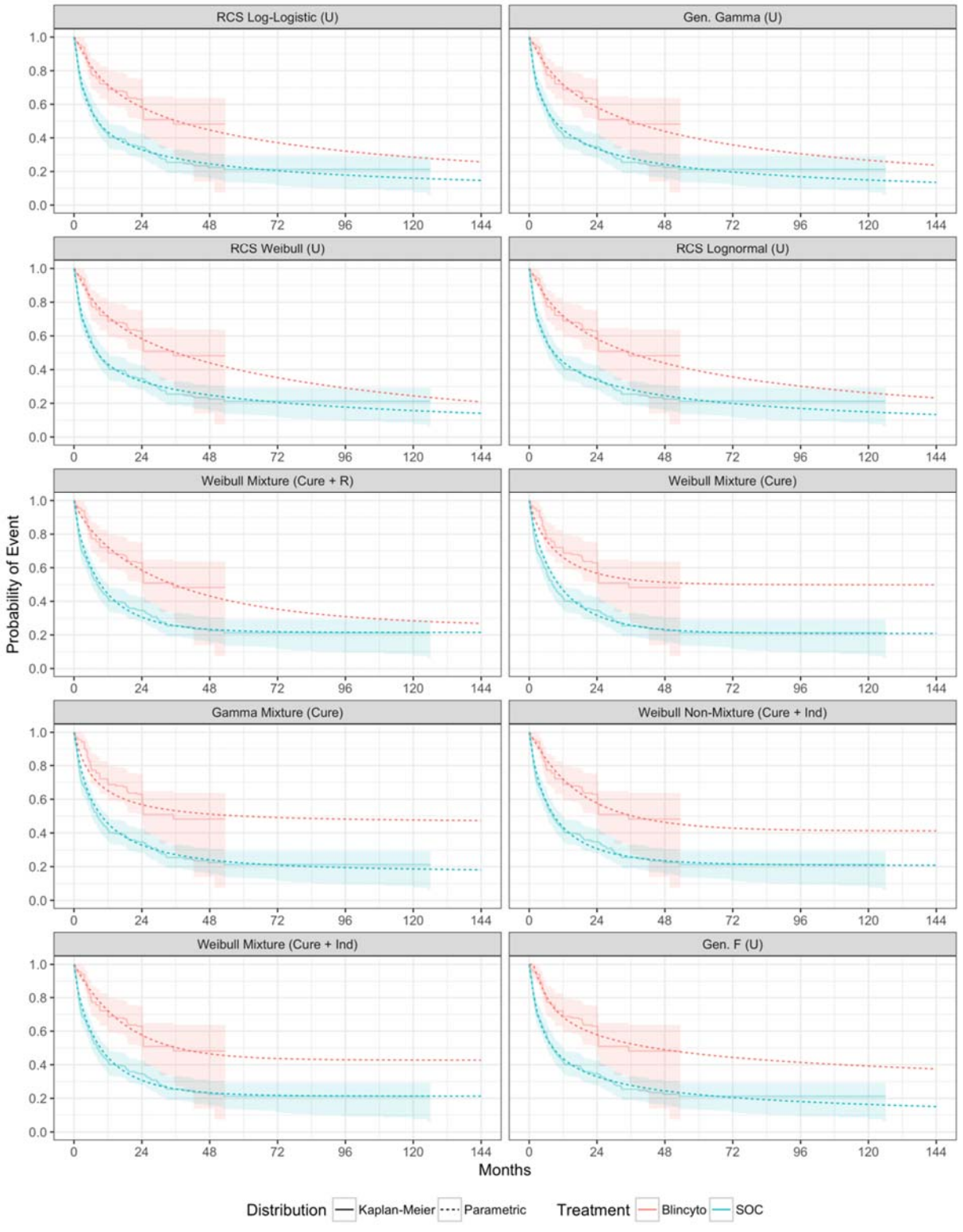


Figure 73. Survival probabilities to 12 years for all parametric survival distributions fit to RFS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATE weights




Lognormal (U)
Weibull Non-Mixture (Cure)
Gompertz (U)
RCS Weibull (R)
Lognormal Mixture (Cure)
Gamma Non-Mixture (Cure)
Weibull Non-Mixture (Cure + R)
Lognormal Mixture (Cure + Ind)
Lognormal Non-Mixture (Cure + Ind)
Log-Logistic (U)
Probability of Event
0.0
0.2
0.4
0.6
0.8
1.0
0
24
48
72
96
120
144
Months
Distribution
Kaplan-Meier
Parametric
Treatment
Blincyto
SOC


RCS Log-Logistic (U)
Gen. Gamma (U)
RCS Weibull (U)
RCS Lognormal (U)
Weibull Mixture (Cure + R)
Weibull Mixture (Cure)
Gamma Mixture (Cure)
Weibull Non-Mixture (Cure + Ind)
Weibull Mixture (Cure + Ind)
Gen. F (U)
Probability of Event
0.0
0.2
0.4
0.6
0.8
1.0
0
24
48
72
96
120
144
Months
Distribution
Kaplan-Meier
Parametric
Treatment
Blincyto
SOC

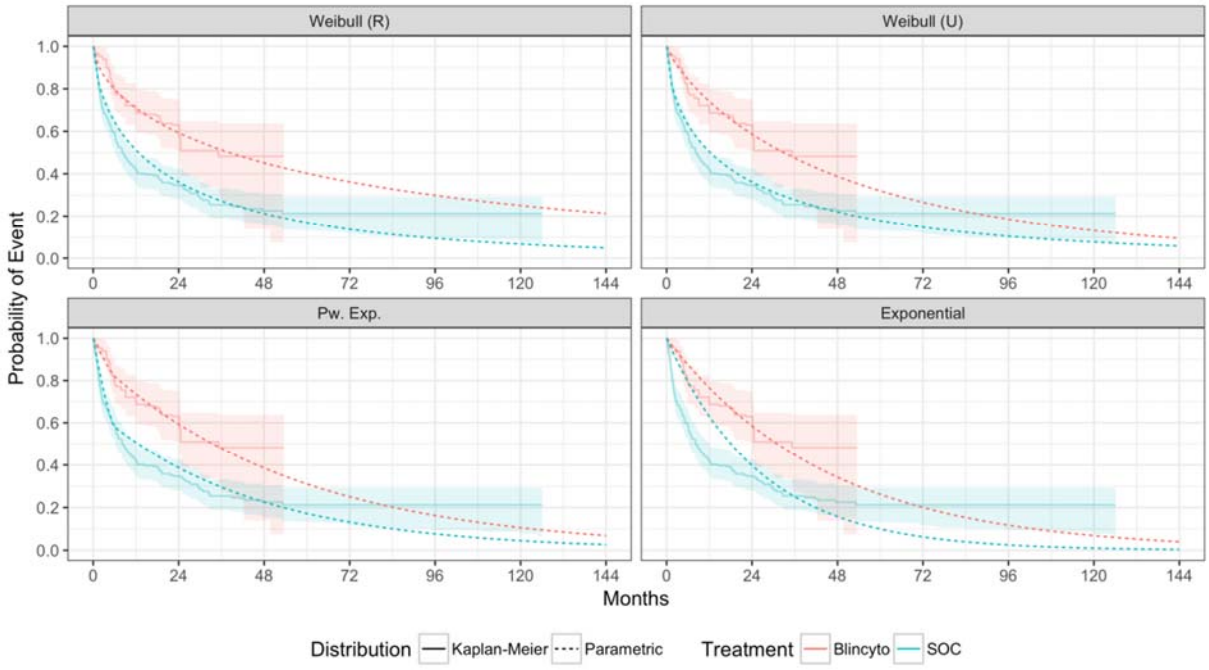
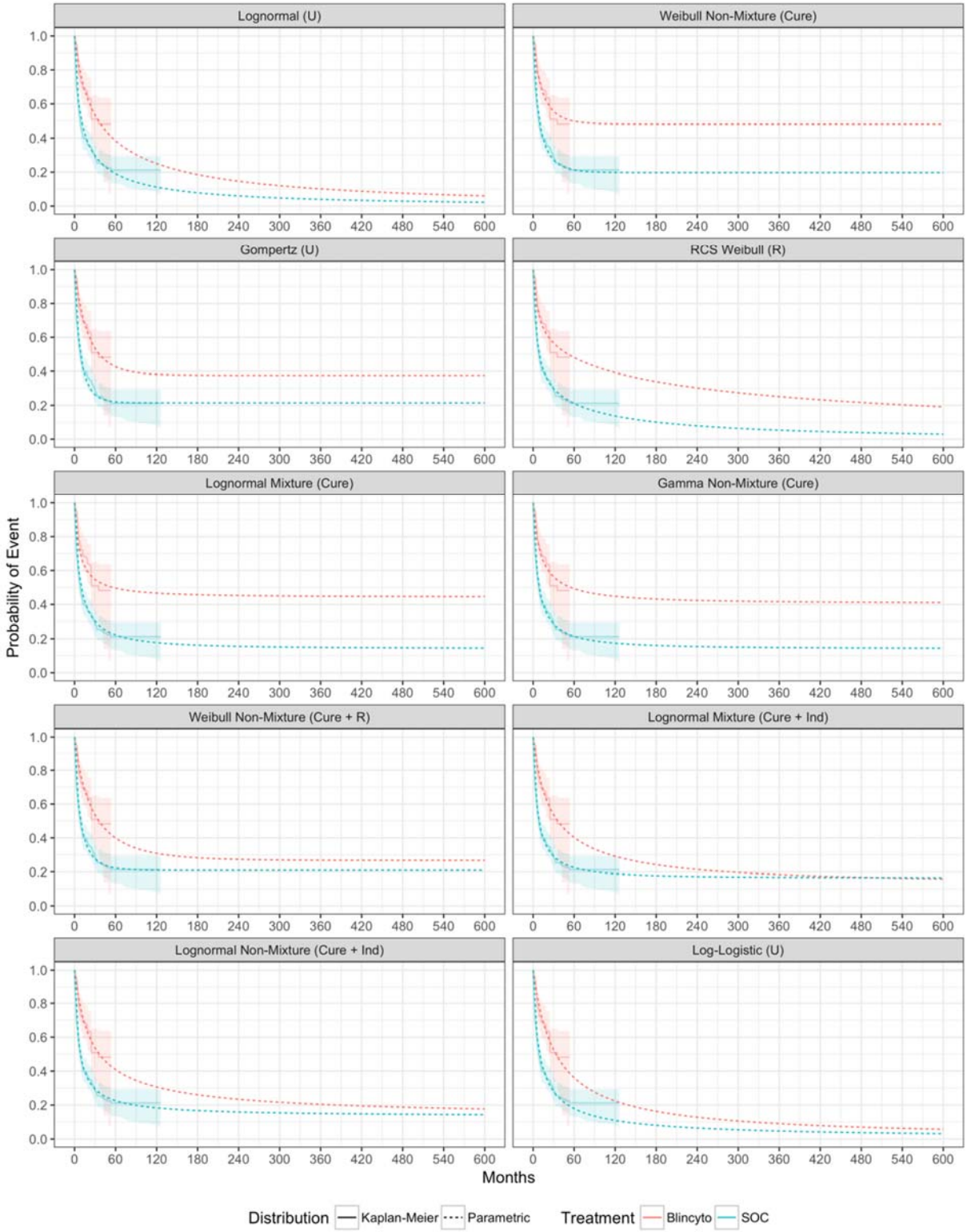
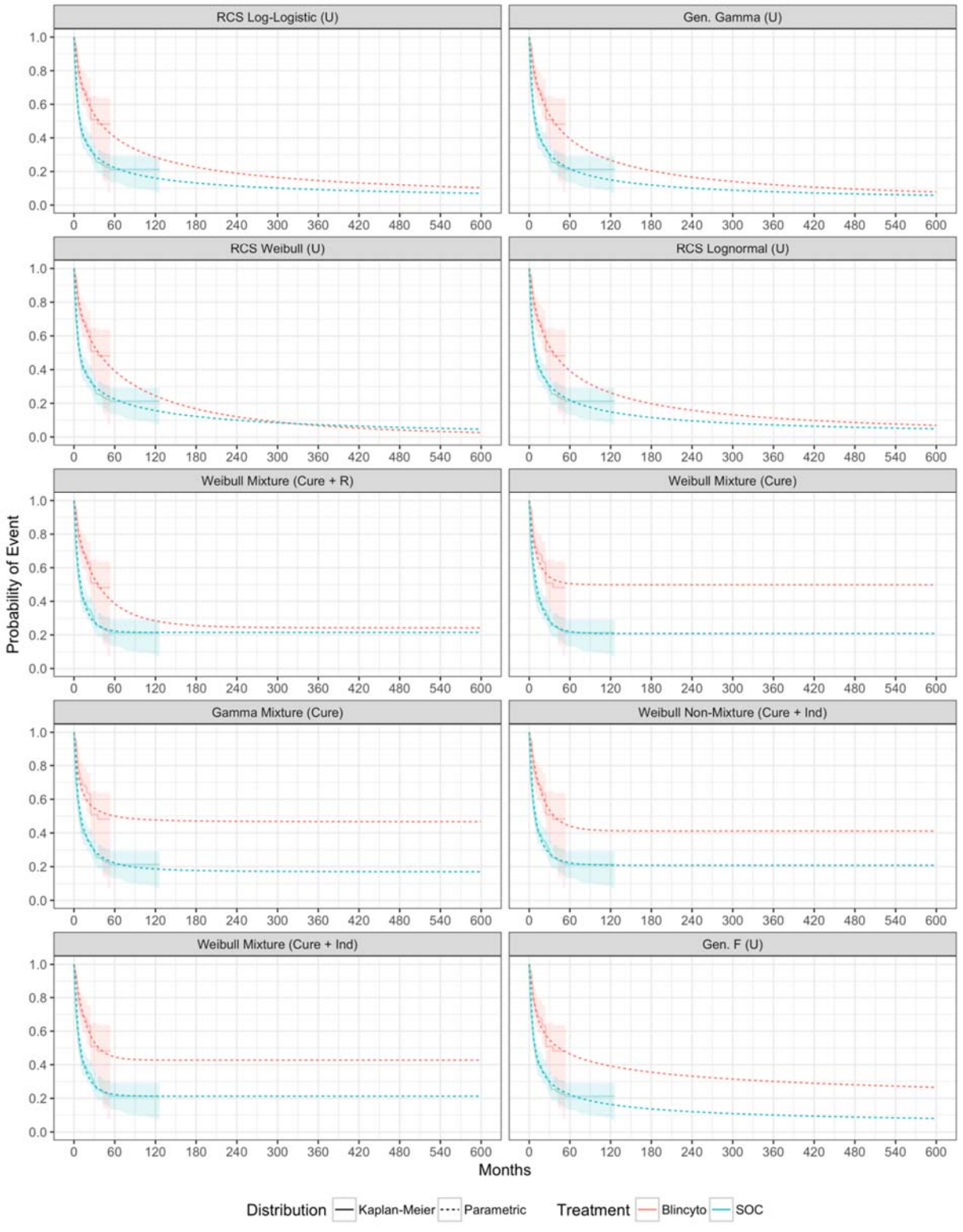

Weibull (R)
Weibull (U)
Pw. Exp.
Exponential
Probability of Event
1.0
0.8
0.6
0.4
0.2
0.0
0
24
48
72
96
120
144
Months
Distribution
Kaplan-Meier
Parametric
Treatment
Blincyto
SOC

Figure 74. Survival probabilities to 50 years for all parametric survival distributions fit to RFS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATE weights




RCS Log-Logistic (U)
Gen. Gamma (U)
RCS Weibull (U)
RCS Lognormal (U)
Weibull Mixture (Cure + R)
Weibull Mixture (Cure)
Gamma Mixture (Cure)
Weibull Non-Mixture (Cure + Ind)
Weibull Mixture (Cure + Ind)
Gen. F (U)
Probability of Event
0.0
0.2
0.4
0.6
0.8
1.0
0
60
120
180
240
300
360
420
480
540
600
Months
Distribution
Kaplan-Meier
Parametric
Treatment
Blincyto
SOC

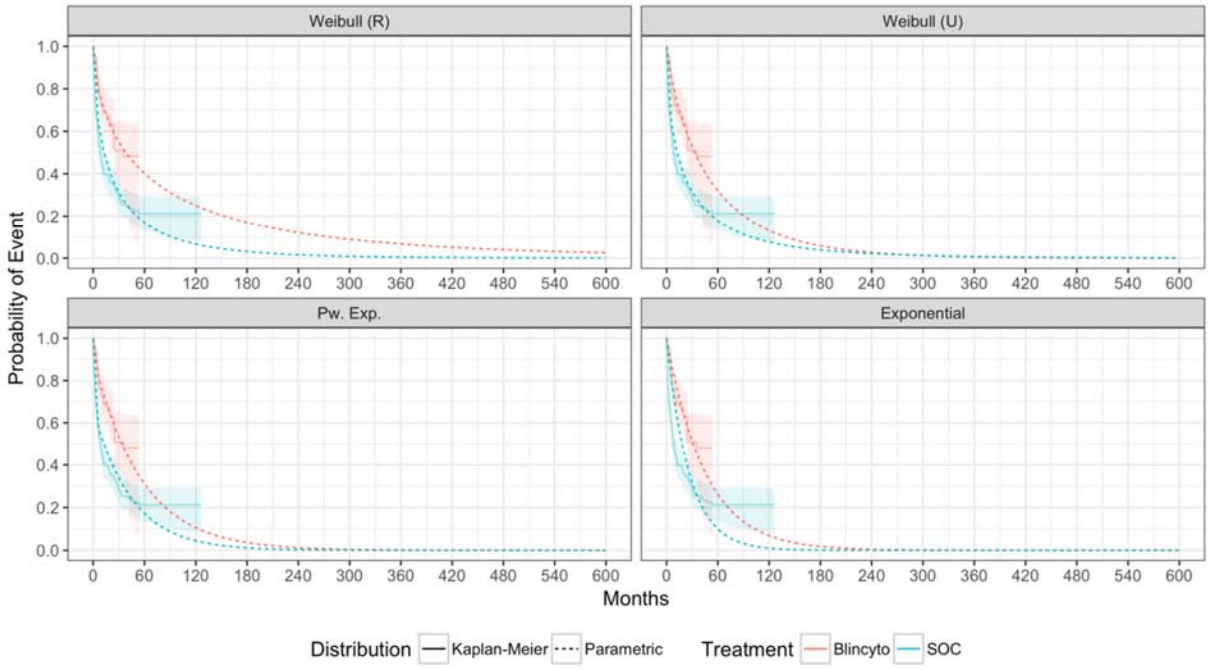


Table 92. Estimated cure fractions for all parametric survival distributions fit to RFS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATE weights

Distribution	Type	Blinatumomab	SOC
Lognormal	Restricted	0.0%	0.0%
Gen. Gamma	Restricted	0.0%	0.0%
RCS Log-Logistic	Restricted	0.0%	0.0%
RCS Lognormal	Restricted	0.0%	0.0%
Gompertz	Restricted	49.9%	20.4%
Lognormal Non-Mixture	Cure	36.0%	11.0%
Lognormal Mixture	Cure + Restricted	19.0%	16.2%
Lognormal Non-Mixture	Cure + Restricted	15.8%	13.4%
Gen. F	Restricted	0.0%	0.0%
Log-Logistic	Restricted	0.0%	0.0%
Lognormal	Unrestricted	0.0%	0.0%
Weibull Non-Mixture	Cure	48.2%	19.7%
Gompertz	Unrestricted	37.6%	21.4%
RCS Weibull	Restricted	0.0%	0.0%
Lognormal Mixture	Cure	44.4%	14.3%
Gamma Non-Mixture	Cure	40.8%	14.2%
Weibull Non-Mixture	Cure + Restricted	26.7%	20.9%
Lognormal Mixture	Cure + Unrestricted	12.9%	16.4%

Lognormal Non-Mixture	Cure + Unrestricted	12.9%	13.5%
Log-Logistic	Unrestricted	0.0%	0.0%
RCS Log-Logistic	Unrestricted	0.0%	0.0%
Gen. Gamma	Unrestricted	0.0%	0.0%
RCS Weibull	Unrestricted	0.0%	0.0%
RCS Lognormal	Unrestricted	0.0%	0.0%
Weibull Mixture	Cure + Restricted	24.2%	21.5%
Weibull Mixture	Cure	49.8%	20.9%
Gamma Mixture	Cure	46.6%	17.0%
Weibull Non-Mixture	Cure + Unrestricted	41.2%	20.8%
Weibull Mixture	Cure + Unrestricted	42.8%	21.3%
Gen. F	Unrestricted	0.0%	0.0%
Weibull	Restricted	0.0%	0.0%
Weibull	Unrestricted	0.0%	0.0%
Pw. Exp.	Unrestricted	0.0%	0.0%
Exponential	Restricted	0.0%	0.0%

1.2.2 Overall Survival

Figure 75. Kaplan Meier estimates of OS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATE weights

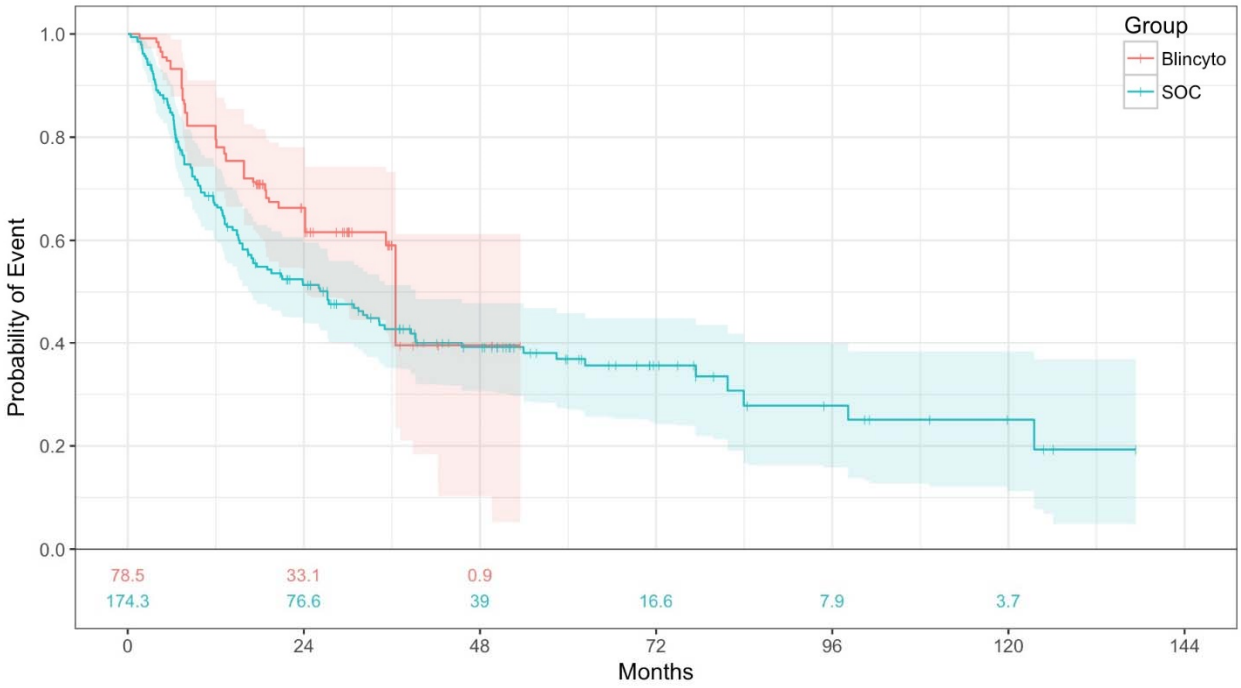


Figure 76. Fit statistics all parametric distributions fit to OS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATE weights

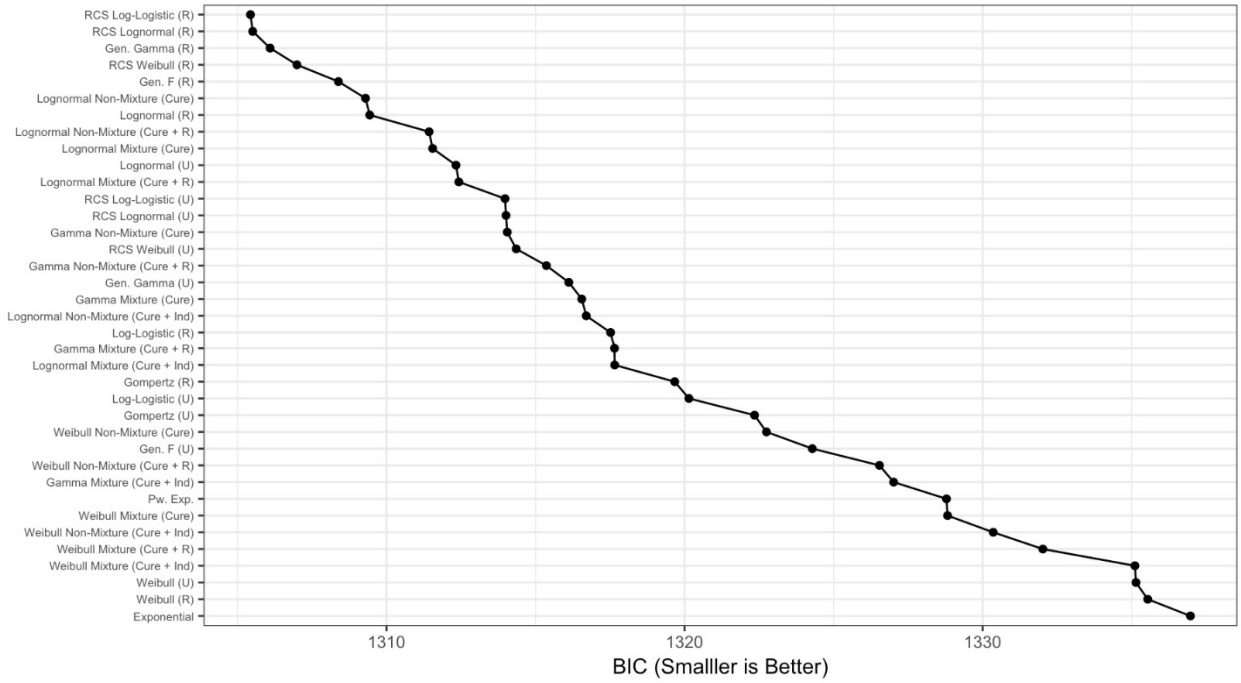


Figure 77. Treatment effect counterfactual plots for OS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATE weights

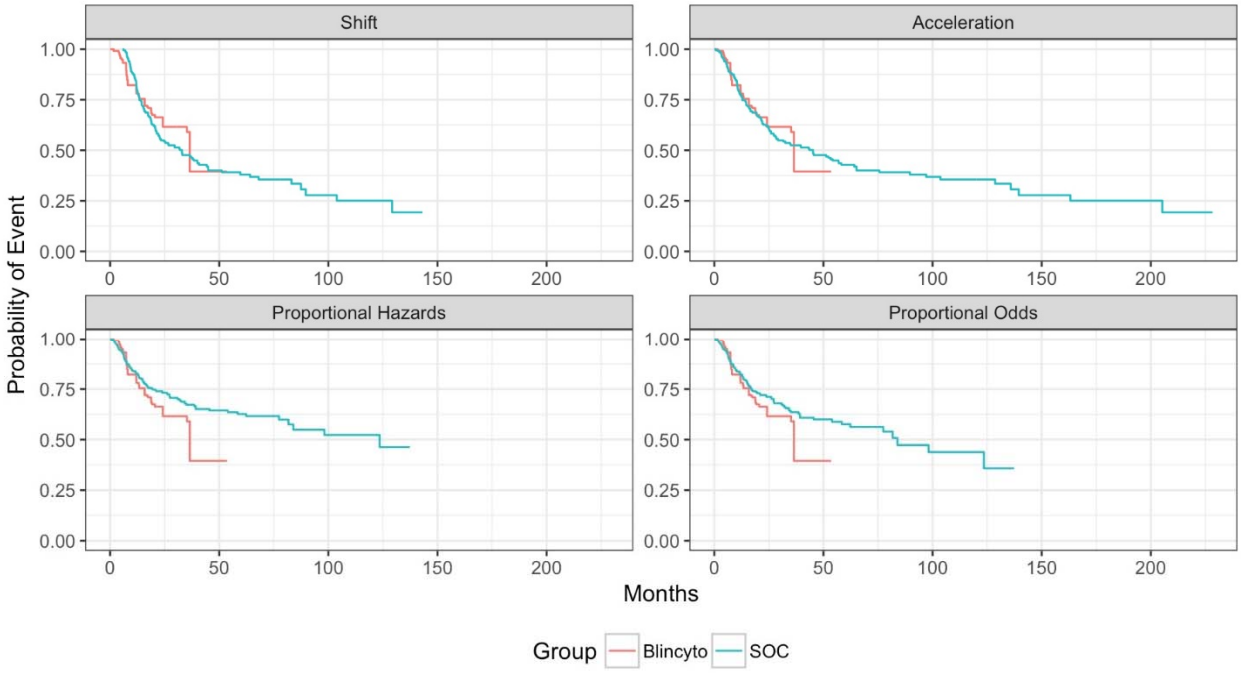


Figure 78. Schoenfeld residuals for OS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATE weights

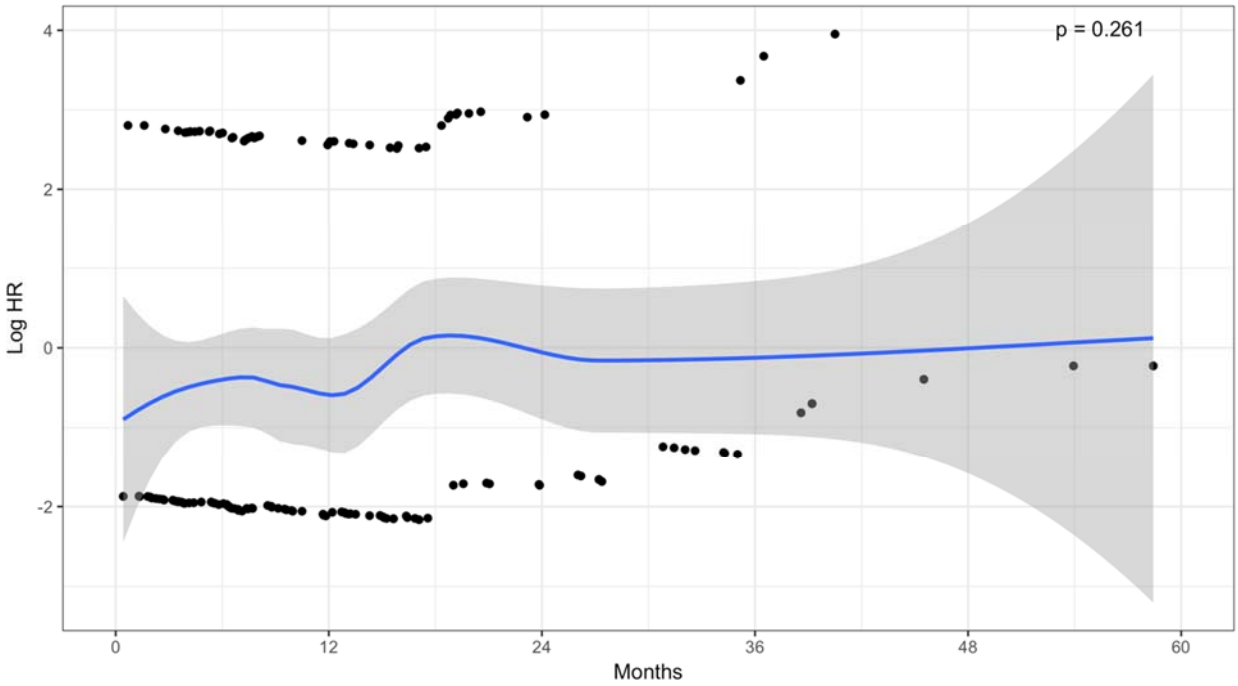
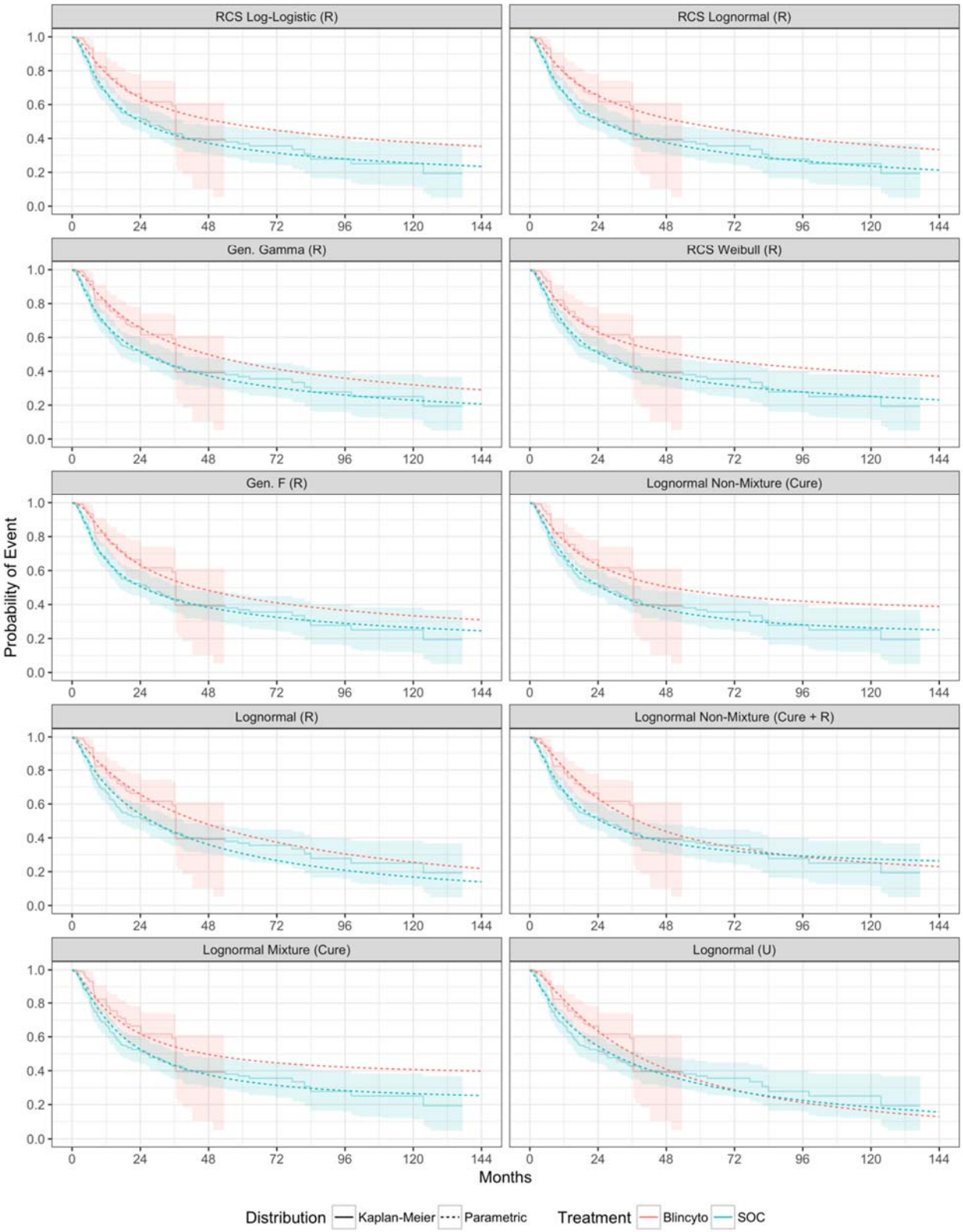
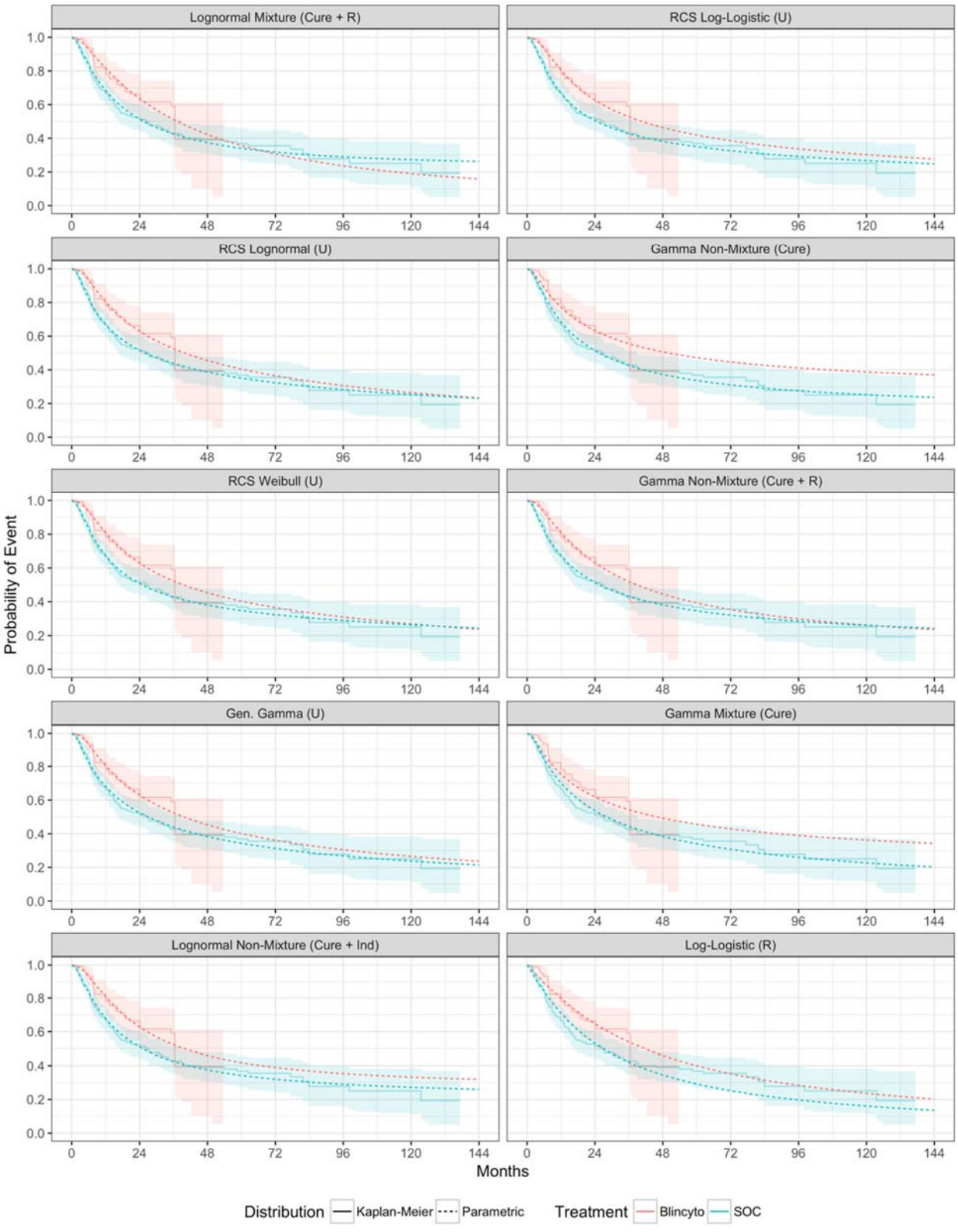
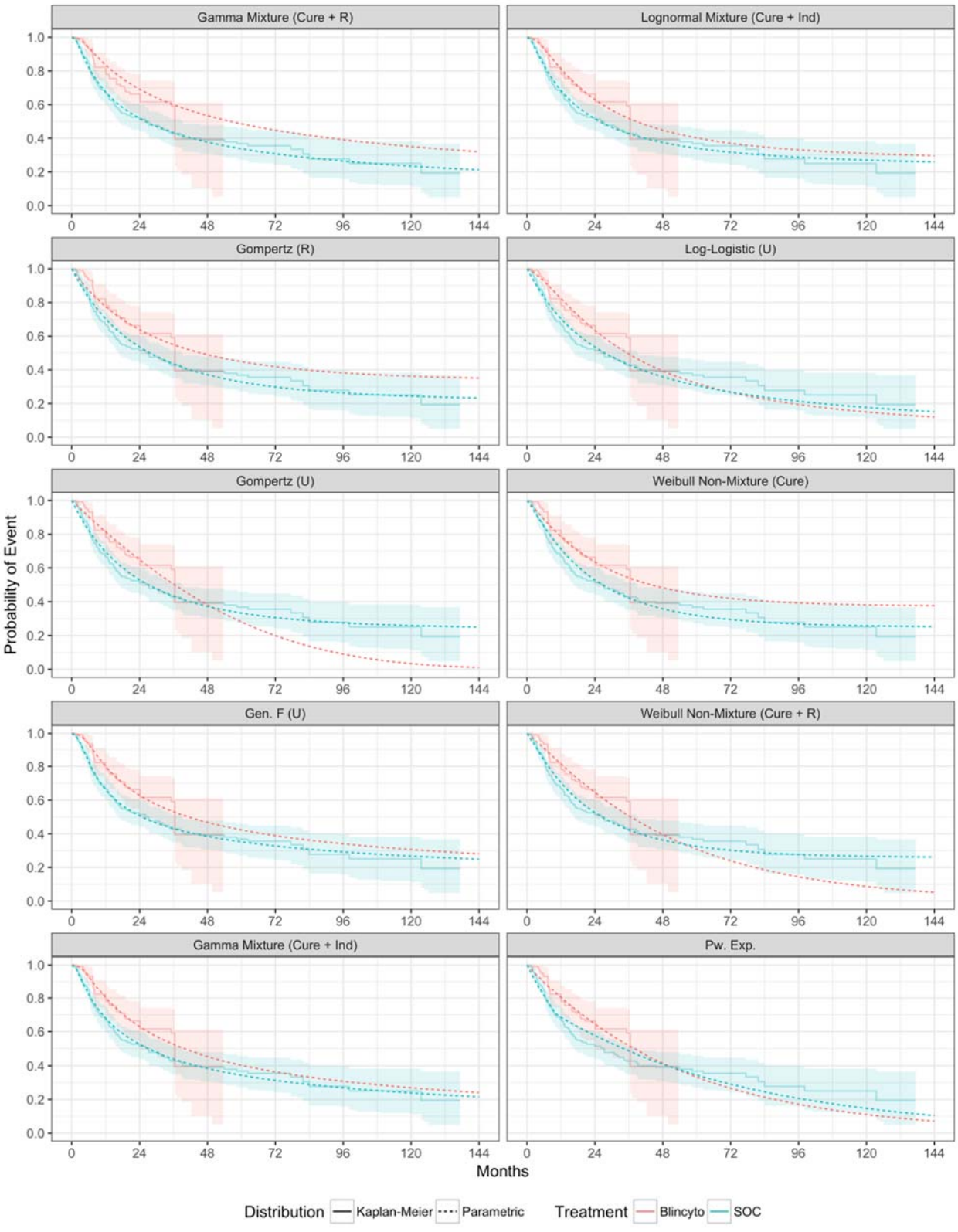


Figure 79. Survival probabilities to 12 years for all parametric distributions fit to OS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATE weights




Lognormal Mixture (Cure + R)
RCS Log-Logistic (U)
RCS Lognormal (U)
Gamma Non-Mixture (Cure)
RCS Weibull (U)
Gamma Non-Mixture (Cure + R)
Gen. Gamma (U)
Gamma Mixture (Cure)
Lognormal Non-Mixture (Cure + Ind)
Log-Logistic (R)
Probability of Event
0.0
0.2
0.4
0.6
0.8
1.0
0
24
48
72
96
120
144
Months
Distribution
Kaplan-Meier
Parametric
Treatment
Blincyto
SOC


Gamma Mixture (Cure + R)
Lognormal Mixture (Cure + Ind)
Gompertz (R)
Log-Logistic (U)
Gompertz (U)
Weibull Non-Mixture (Cure)
Gen. F (U)
Weibull Non-Mixture (Cure + R)
Gamma Mixture (Cure + Ind)
Pw. Exp.
Probability of Event
0.0
0.2
0.4
0.6
0.8
1.0
0
24
48
72
96
120
144
Months
Distribution
Kaplan-Meier
Parametric
Treatment
Blincyto
SOC

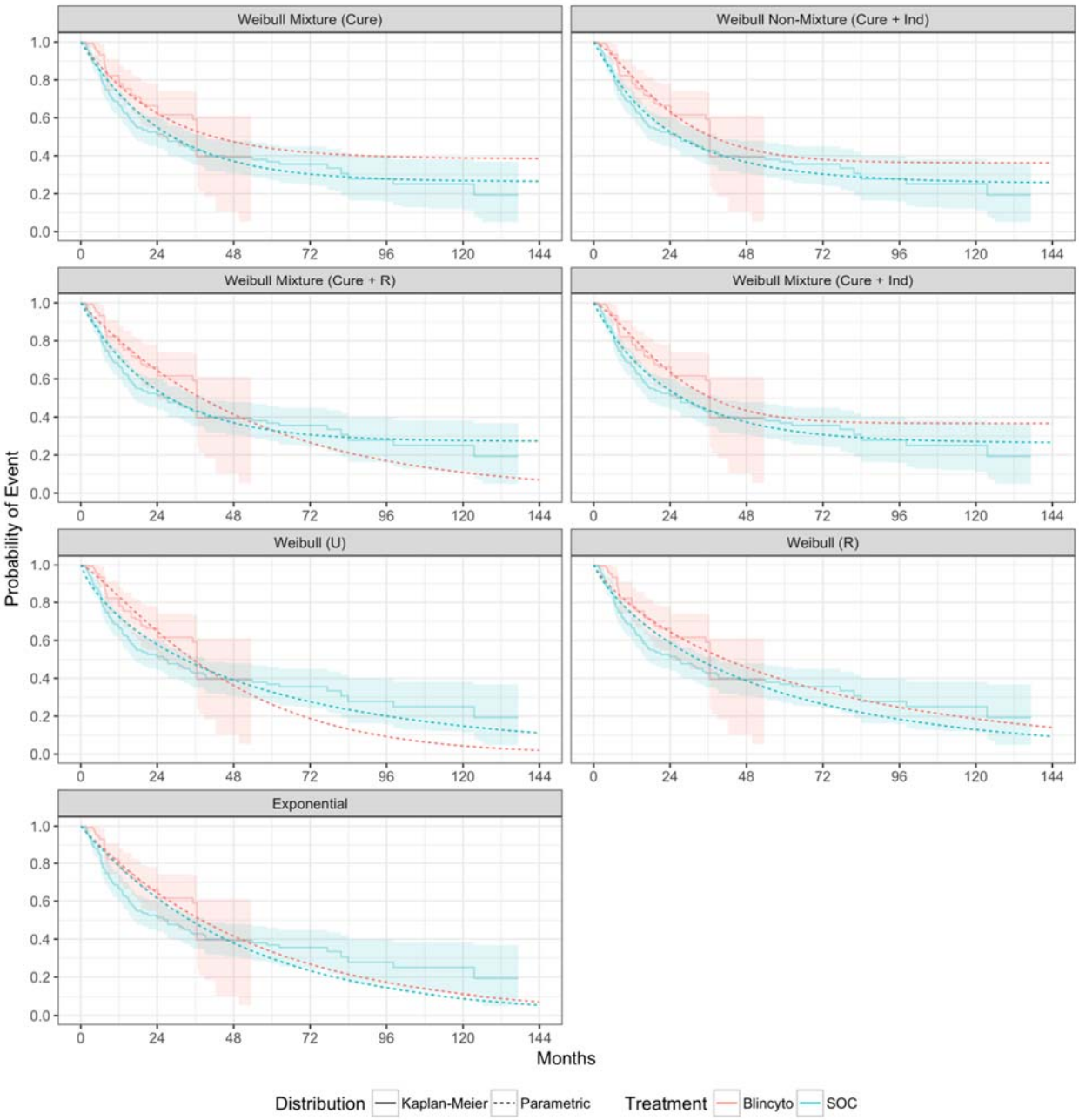
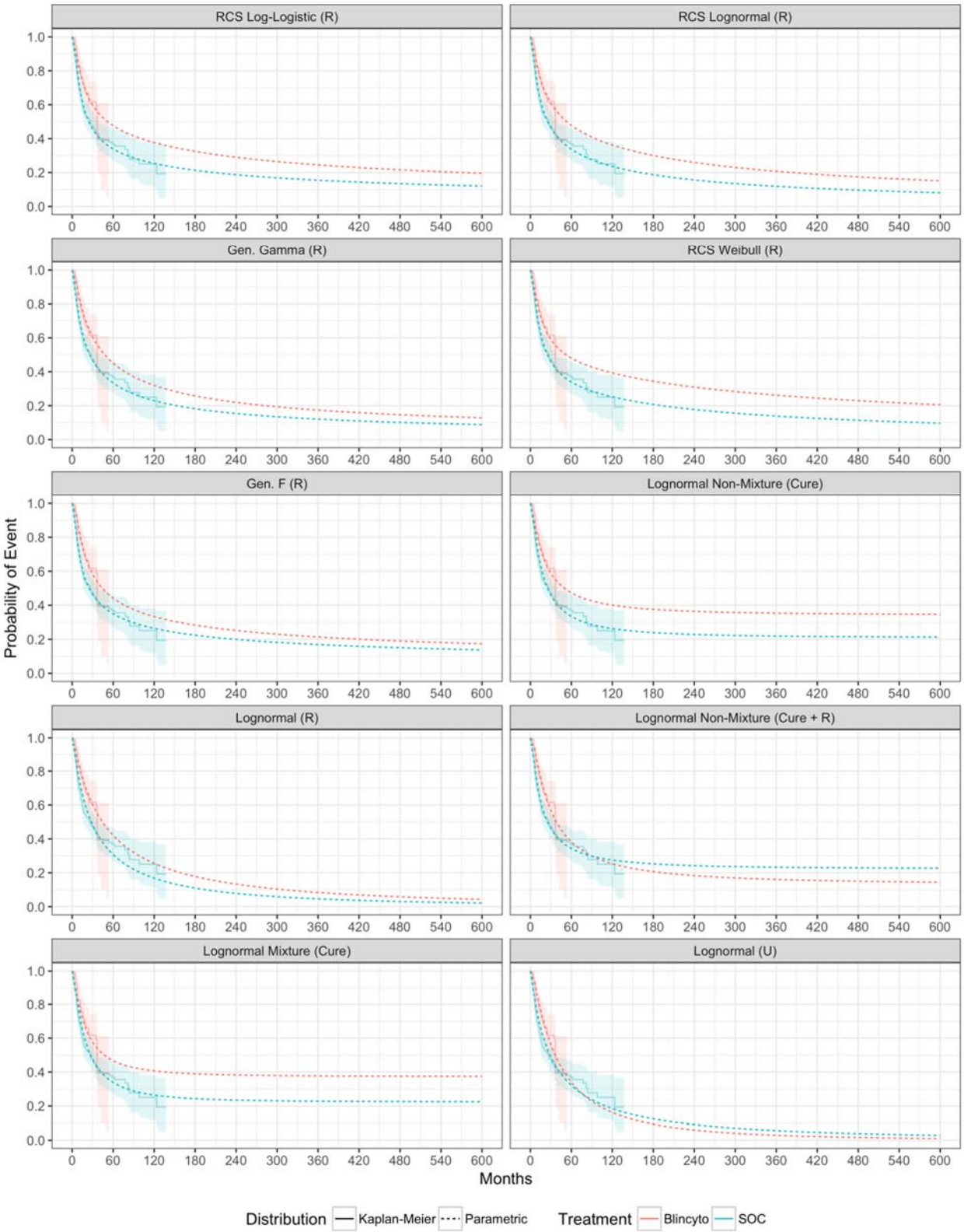
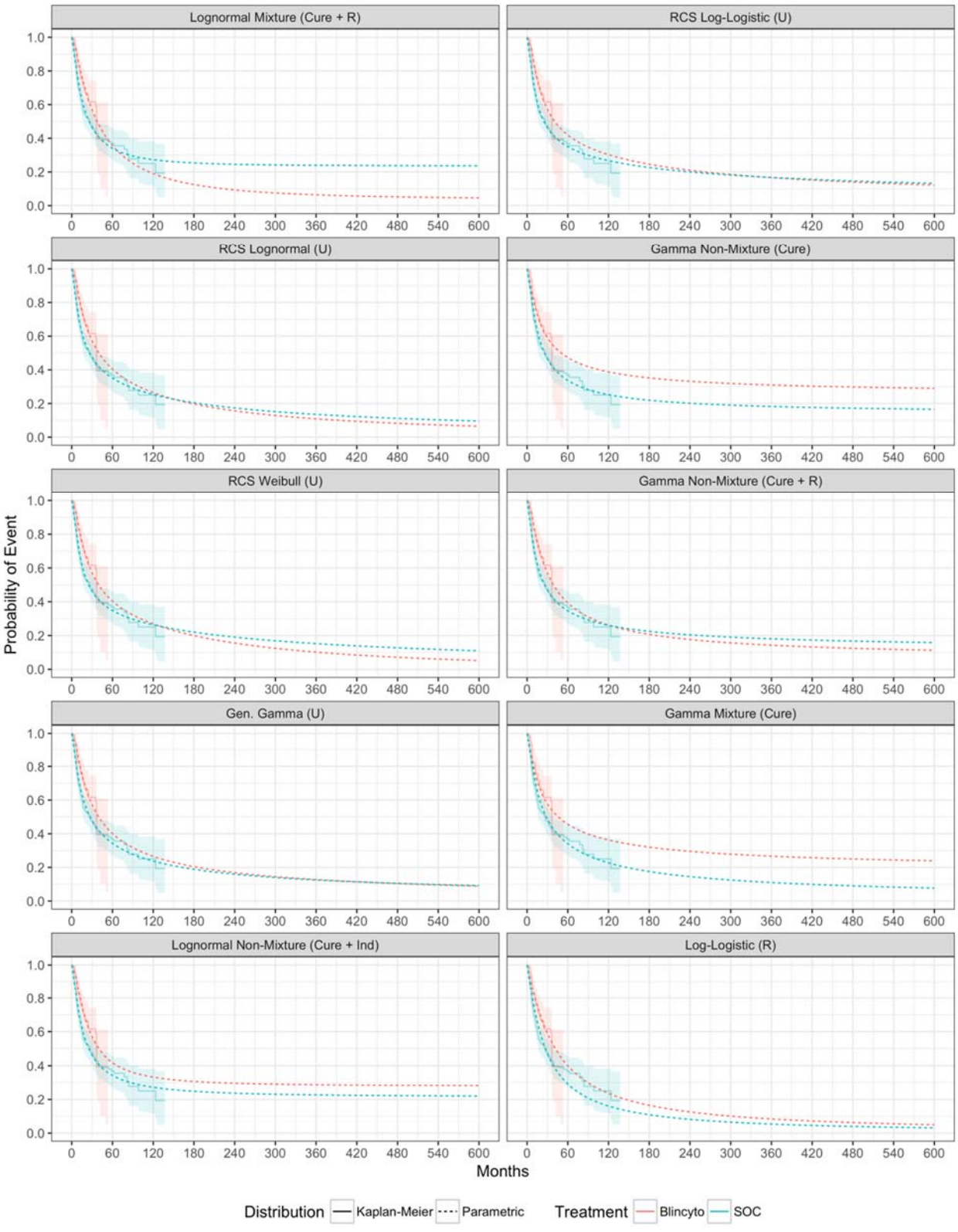
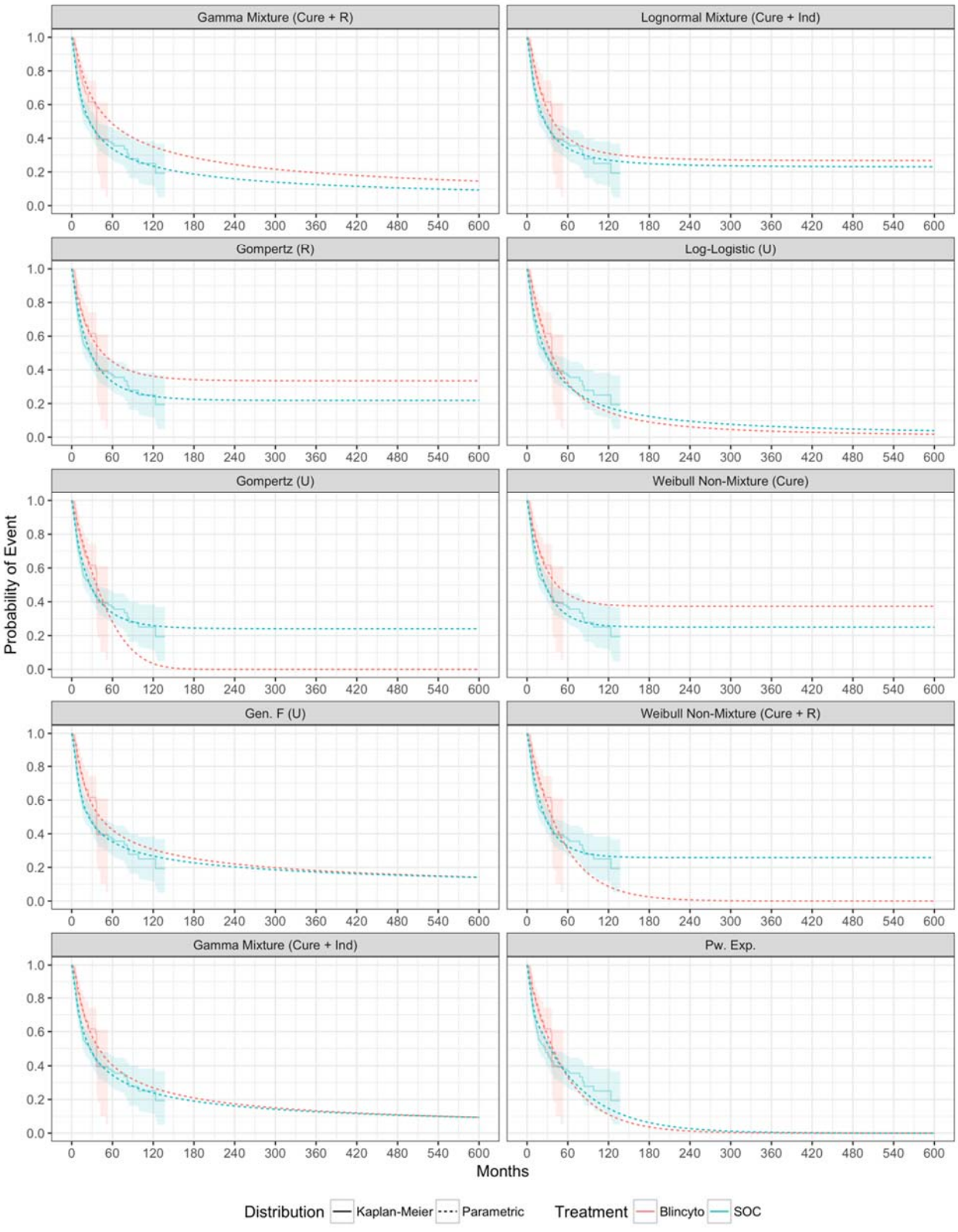

Weibull Mixture (Cure)
Weibull Non-Mixture (Cure + Ind)
Weibull Mixture (Cure + R)
Weibull Mixture (Cure + Ind)
Weibull (U)
Weibull (R)
Exponential
Probability of Event
Months
0.0
0.2
0.4
0.6
0.8
1.0
0
24
48
72
96
120
144
Distribution
Kaplan-Meier
Parametric
Treatment
Blincyto
SOC

Figure 80. Survival probabilities to 50 years for all parametric distributions fit to OS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATE weights




Lognormal Mixture (Cure + R)
RCS Log-Logistic (U)
RCS Lognormal (U)
Gamma Non-Mixture (Cure)
RCS Weibull (U)
Gamma Non-Mixture (Cure + R)
Gen. Gamma (U)
Gamma Mixture (Cure)
Lognormal Non-Mixture (Cure + Ind)
Log-Logistic (R)
Probability of Event
0.0
0.2
0.4
0.6
0.8
1.0
0
60
120
180
240
300
360
420
480
540
600
Months
Distribution
Kaplan-Meier
Parametric
Treatment
Blincyto
SOC


Gamma Mixture (Cure + R)
Lognormal Mixture (Cure + Ind)
Gompertz (R)
Log-Logistic (U)
Gompertz (U)
Weibull Non-Mixture (Cure)
Gen. F (U)
Weibull Non-Mixture (Cure + R)
Gamma Mixture (Cure + Ind)
Pw. Exp.
Probability of Event
0.0
0.2
0.4
0.6
0.8
1.0
0
60
120
180
240
300
360
420
480
540
600
Months
Distribution
Kaplan-Meier
Parametric
Treatment
Blincyto
SOC

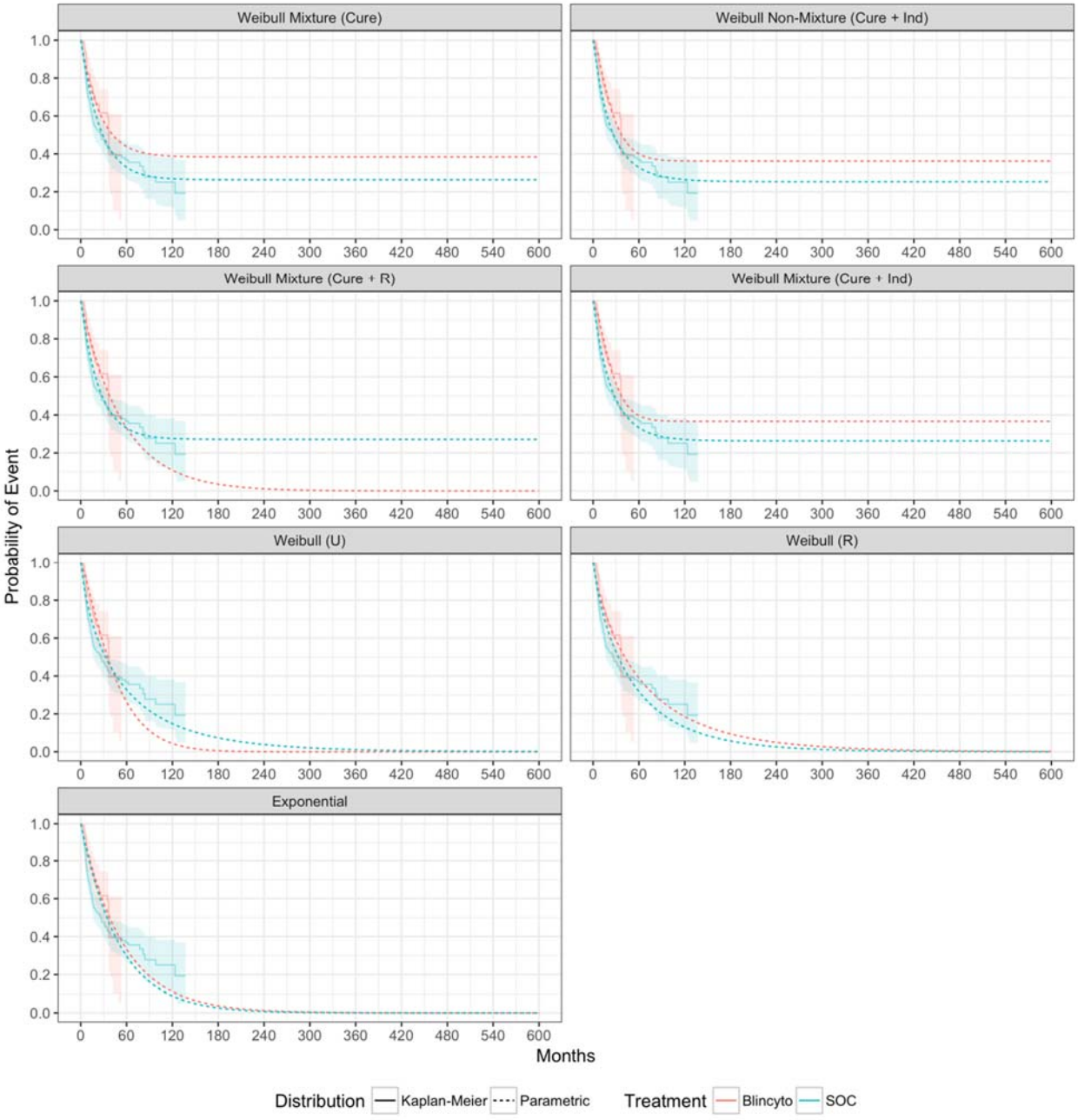


Table 93. Estimated cure fractions for all parametric distributions fit to OS for patients receiving blinatumomab and SOC from propensity matched analysis of BLAST and historical control study, ATE weights

Distribution	Type	Blinatumomab	SOC
RCS Log-Logistic	Restricted	0.0%	0.0%
RCS Lognormal	Restricted	0.0%	0.0%
Gen. Gamma	Restricted	0.0%	0.0%
RCS Weibull	Restricted	0.0%	0.0%
Gen. F	Restricted	0.1%	0.0%
Lognormal Non-Mixture	Cure	34.3%	20.9%
Lognormal	Restricted	0.0%	0.0%
Lognormal Non-Mixture	Cure + Restricted	13.1%	22.3%

Lognormal Mixture	Cure	37.4%	22.4%
Lognormal	Unrestricted	0.0%	0.0%
Lognormal Mixture	Cure + Restricted	3.6%	23.5%
RCS Log-Logistic	Unrestricted	0.0%	0.0%
RCS Lognormal	Unrestricted	0.0%	0.0%
Gamma Non-Mixture	Cure	25.4%	13.7%
RCS Weibull	Unrestricted	0.0%	0.0%
Gamma Non-Mixture	Cure + Restricted	4.2%	9.7%
Gen. Gamma	Unrestricted	0.0%	0.0%
Gamma Mixture	Cure	17.6%	0.0%
Lognormal Non-Mixture	Cure + Unrestricted	28.1%	21.6%
Log-Logistic	Restricted	0.0%	0.0%
Gamma Mixture	Cure + Restricted	0.0%	0.0%
Lognormal Mixture	Cure + Unrestricted	26.7%	22.9%
Gompertz	Restricted	33.5%	21.9%
Log-Logistic	Unrestricted	0.0%	0.0%
Gompertz	Unrestricted	0.0%	24.0%
Weibull Non-Mixture	Cure	37.3%	25.0%
Gen. F	Unrestricted	0.0%	0.0%
Weibull Non-Mixture	Cure + Restricted	0.0%	25.8%
Gamma Mixture	Cure + Unrestricted	0.0%	0.0%
Pw. Exp.	Unrestricted	0.0%	0.0%
Weibull Mixture	Cure	38.4%	26.4%
Weibull Non-Mixture	Cure + Unrestricted	36.2%	25.3%
Weibull Mixture	Cure + Restricted	0.0%	27.1%
Weibull Mixture	Cure + Unrestricted	36.6%	26.3%
Weibull	Unrestricted	0.0%	0.0%
Weibull	Restricted	0.0%	0.0%
Exponential	Restricted	0.0%	0.0%

Figure 10. RFS *without* censoring at allogeneic HSCT and post-BLINCYTO chemotherapy

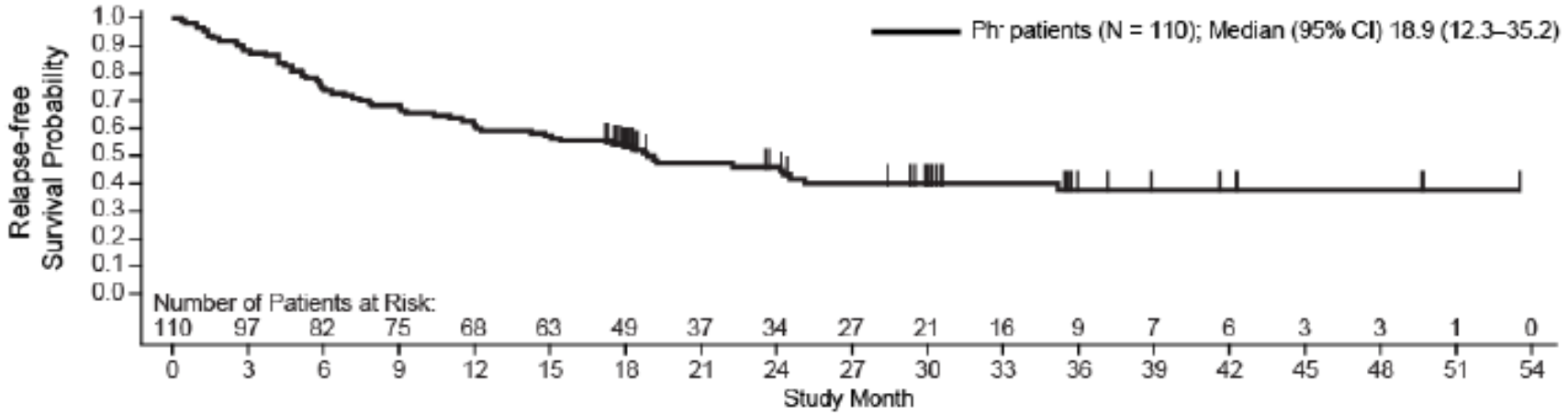
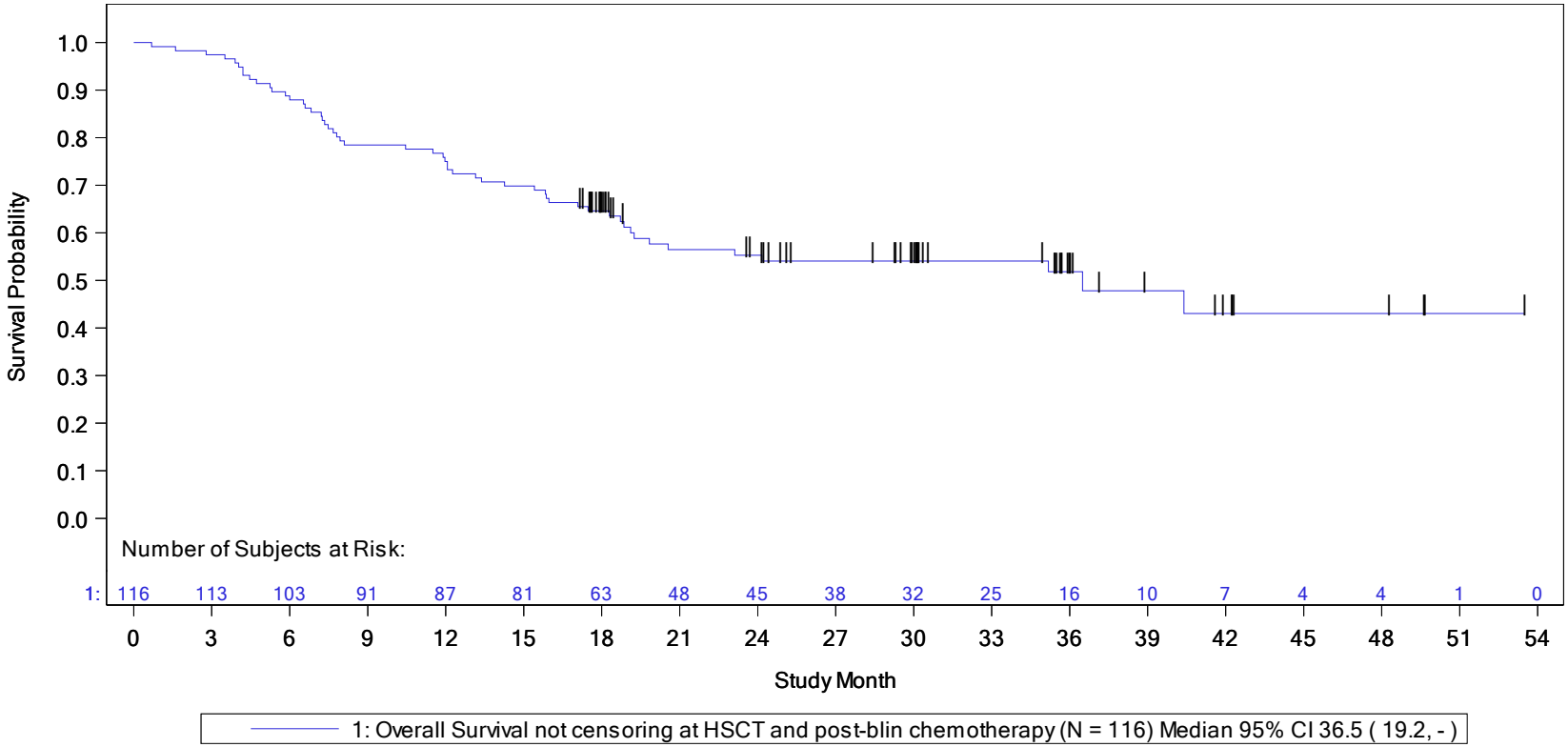


Figure 13. OS in BLAST *without* censoring at allogeneic HSCT and post-BLINCYTO chemotherapy



Censor indicated by Vertical Bar

Program: /userdata/stat/amg103/meta/mrd_2017/analysis/reg_quest/figures/program/f-odac-os.sas
Output: fodac-12-odac-os-fasfl.rtf (Date Generated: 08FEB18 00:31) Source Data: a203.adsl, a203.adtteeff

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

Single Technology Appraisal (STA)

Blinatumomab for acute lymphoblastic leukaemia for people with minimal residual disease activity in remission [ID1036]

Dear Gavin and Kawitha,

The Evidence Review Group, ScHARR, and the technical team at NICE have looked at the submission received on 31 October 2017 from Amgen. In general they felt that it is well presented and clear. However, the ERG and the NICE technical team would like further clarification on the clinical and cost effectiveness data (see questions listed at end of letter).

The ERG and the technical team at NICE will be addressing these issues in their reports.

Please provide your written response to the clarification questions by **5pm** on **7 December 2017**. Your response and any supporting documents should be uploaded to NICE Docs: https://appraisals.nice.org.uk/request/38852

Two versions of this written response should be submitted; one with academic/commercial in confidence information clearly marked and one from which this information is removed.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, and all information submitted under 'academic in confidence' in yellow.

If you present data that is not already referenced in the main body of your submission and that data is seen to be academic/commercial in confidence information, please complete the attached checklist for in confidence information.

Please do not 'embed' documents (i.e. PDFs, spreadsheets) within your response as this may result in your information being displaced or unreadable.

If you have any queries on the technical issues raised in this letter, please contact Sana Khan, Technical Lead (Sana.Khan@nice.org.uk). Any procedural questions should be addressed to Thomas Feist, Project Manager TACommA@nice.org.uk.

Yours sincerely

Eleanor Donegan
Technical Advisor – Appraisals
Centre for Health Technology Evaluation

Encl. checklist for in confidence information

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

Section A: Clarification on effectiveness data

Scope/General

A1. PRIORITY. Please provide the proposed wording of the marketing authorisation for blinatumomab for the indication under appraisal.

A2. PRIORITY. Company submission (CS), Table 1 decision problem, page 15. The CS states "*Although the cost effectiveness evidence does not consider the Ph+ population or later remission states, due to the substantial unmet need across all sub-populations, blinatumomab should be considered for use in alignment with its full anticipated marketing authorisation."*

- Please comment on the expected clinical effectiveness and cost-effectiveness of blinatumomab for patients in second or subsequent haematological CR.
- Please comment on the expected clinical effectiveness and cost-effectiveness of blinatumomab for patients with Ph+ disease.

A3. PRIORITY. Please clarify why the dosing regimen in the licence authorisation is 28µg/day, given that the dosage employed in BLAST was 15µg/m^2.

A4. Please provide a PDF file of the following reference (reference 5 in CS)
Gökbuget N, Dombret, H., Bonifacio, M., Reichle, A., Graux, C., Havelange, V., Buss, E. C., Faul, C., Bruggemann, M., Ganser, A., Stieglmaier, J., Wessels, H., Haddad, V., Zugmaier, G., Nagorsen, D., & Bargou, R. C. BLAST: A Confirmatory, Single-Arm, Phase 2 Study of Blinatumomab, a Bispecific T-Cell Engager (BiTE®) Antibody Construct, in Patients with Minimal Residual Disease B-Precursor Acute Lymphoblastic Leukemia (ALL). Blood 2014;124:379.

Literature searching

A5. Company submission, Appendix D Identification, selection and synthesis of clinical evidence:

- The searches were last conducted on 19th May 2017. Please update the search and confirm that no relevant studies have been found since then.
- Please state the exact dates and terms used to search for conference proceedings.
- Table 73, page 186: Other limits/considerations used in the study selection process across the clinical efficacy/safety SLR suggests that registries were searched. If trial registries were searched, please state the source including search strategies. Please give reasons if trial registries were not searched.
- Given that no adverse event studies were identified for blinatumomab for treatment of acute lymphoblastic leukaemia (ALL) with minimal residual disease from the searches presented in Appendix C, please explain why separate adverse events searches were not carried out for blinatumomab only (for information, searches for blinatumomab gives 233 records in PubMed alone).

Systematic review process

A6. Company submission, Appendix D, page 183. Please clarify why three articles were excluded (Figure 46, page 188) for not evaluating treatment of interest, when the inclusion

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

criteria (Table 72, page 185) does not exclude any interventions. Also, given that there is so little evidence available for blinatumomab, please explain why studies with less than 10 patients were excluded.

A7. Company submission, Appendix D, page 183. Please confirm how many reviewers conducted data extraction and quality assessment?

Comparative effectiveness

A8 PRIORITY. Company submission, section B.2.9.4, page 70. With respect to the inverse probability of treatment weighting (IPTW) approach used for propensity score adjustment:

- Please justify why this method has been used over other possible methods that use observational data to inform treatment effectiveness.
- Please clarify how the assumption of selection on observables was assessed.
- Please clarify how the assumption of overlap between the two studies was assessed. Is this considered to be justified for average treatment effect (ATE) and average treatment effect on the treated (ATT) weights?
- A subgroup of BLAST (n=73), was used to estimate comparative effectiveness. Please clarify why each of the listed criteria was required (as opposed to using the whole BLAST trial population).
- Please clarify whether the assumption of ignorability of treatment is considered to be justified for ATE and ATT weights.

A9 PRIORITY. Company submission, section B.2.9.4, page 71. Candidate variables for the propensity score model are provided. Please also provide the final used model.

A10. Company submission, figures 17 and 18, pages 74 and 75. The included figures indicate time from randomisation to event. Please clarify whether the figures and related statistics relate to time from first blinatumomab treatment for BLAST patients and 14 days after the MRD baseline date for patients from the historical comparator study 20120148.

A11 PRIORITY. Company submission, appendix L, pages 216. The start date for time to event outcomes was defined as "*14 days after the MRD baseline date for 20120148 patients and the date of the first blinatumomab treatment for BLAST patients*"

- Please justify why the median time from MRD detection to first blinatumomab dose is the most appropriate cut-point to use
- Please provide further statistics for this variable e.g. what was the range of values observed for the time between MRD detection and first blinatumomab dose?

How many RFS/OS events occurred during the first 14 days, leading to exclusion from the historical control data set?

A12. Please provide details of prior ALL treatments received by patients in the BLAST study.

A13. Company submission, page 67. Please clarify whether the historical control study included patients treated in years 2000-2017 (as suggested in the CS), or 2000-2014 (as suggested in the Observational Study Report).

A14. Company submission, section B.2.9.3, page 68. Please clarify why Ph+ patients and CR2+ patients were not included in the historical control study.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

A15. Company submission, section B.2.9.3, page 68. Please provide details of any peer reviewed published data from the historical control study. If none are available, please provide further details regarding the study design and enrolment procedures.

A16. Company submission, Appendix D, Figure 47, page 195. Please modify the figure to include the reasons for not enrolling 95 patients out of the total of 211 patients who were screened.

Section B: Clarification on cost-effectiveness data

Literature searching

B1. Company submission, appendix D, page 183. Published cost effectiveness studies. Given that no cost-effectiveness studies were identified for ALL with minimal residual disease from the searches contained in Appendix C, please explain why separate cost-effectiveness searches were not carried out for ALL only i.e. ALL terms combined with an economic evaluations study design filter.

B2. Company submission, appendix D, Table 71, page 185: York Centre for Reviews and Dissemination search algorithm – page 183 of appendix D states that "these databases were searched via the EMBASE®, PubMed, Cochrane Library and York Centre for Reviews and Dissemination". Table 71 shows ALL terms combined with costs/economic terms, but this is not consistent across the other relevant databases i.e. EMBASE®, PubMed, Cochrane Library where it is ALL+MRD and not ALL+ costs/economic terms. Please comment on the likelihood that no economic evaluations have been missed as a consequence of this approach.

B3. Company submission, appendix E, page 198. HRQoL studies. Given that no HRQoL studies were identified for ALL with minimal residual disease from the searches contained in Appendix C, please explain why a separate HRQoL search was not carried out for ALL only i.e. ALL terms combined with HRQoL terms?

Survival modelling

B4. PRIORITY. Company submission, section B.3.3. pages 93-116. Five model types have been applied within the CS (Unrestricted, Restricted, Restricted Cubic Spline, Mixture cure, and Non-Mixture cure). For the general case, and the best fitting model in each category (e.g. Gompertz for unrestricted)

- Please provide mathematical equations for the model. For example, for the Gompertz unrestricted model, please provide the regression equation "*including treatment -group interaction terms for every distributional parameter*".

Please provide the code and output showing how these models have been fitted in the relevant software package (i.e. flexsurv in R for the parametric models, STATA for the cure models)

B5. PRIORITY. Company submission, section B.3.3, pages 93-116. Please clarify why the lognormal mixture cure model (and presumably every other mixture cure model fitted) does not include the expected mortality for the cured fraction (see Lambert, The Stata Journal, 2007, 7(3) i.e. the source of the STATA procedures cited in the CS). Please also comment on the potential bias associated with estimating survival probabilities for cured patients based on a fixed model start age rather than the observed distribution of patient age within the clinical studies.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

B6. PRIORITY. Company submission, section B.3.3. pages 93-116. Please explain the conceptual implications of selecting a cure model for OS but not for RFS. How can patients who are cured in terms of OS still experience relapse according to RFS?

B7. PRIORITY. CS, Section B.3.3. pages 93-116. Please explain whether and how clinical judgement was used to inform the selection of the parametric functions for RFS and OS in the model.

B8. PRIORITY. Company submission, section B.3.3. pages 93-116. Please provide complete sensitivity analyses which include all combinations of functional forms for RFS and OS (assuming same curve type for each treatment group). Please ensure that this analysis does not exclude combinations of OS and RFS where the RFS and OS curves cross (in such instances, please constrain RFS to the minimum of RFS and OS). Please report the results of these analyses in the form of a table including incremental QALYs, incremental costs and ICERs for each comparison.

B9. Company submission, section B.3.3.1, Figure 23 (page 98), figure 30(page 104), appendix P, figure 58 and 64. Please provide model fit statistics (BIC) in a table rather than a line graph.

B10. Company submission, section B.3.3.1, Table 35, page 94. The CS refers to "unrestricted" parametric models which the ERG understands to mean fitting separate models to each treatment group without the inclusion of a treatment effect parameter (a hazard ratio or a constant acceleration factor), or otherwise relaxing this restriction as implemented in the CS through the inclusion of the interaction terms. However, Table 35 includes a tick mark for an HR or AF in the "treatment effect" column for unrestricted models. Please clarify.

B11. Company submission, section B.3.3.1, page 95. Please clarify why 8% of people receiving standard of care (SoC) are assumed to achieve MRD response, given that the clinicians consulted suggested this was "no greater than 10%".

B12. Company submission, section B.3.3.1, page 95-97. In 4 sentences or less, and using non-technical language, please explain what the analysis using the Berry external data mean.

B13. Company submission, section B.3.3.1, page 95-97. Please provide a comparison of the matched OS data to the Berry data, in the same way as was presented for the RFS data (see page 97). These data are discussed but not presented in the CS.

B14. Company submission, section B.3.3. pages 93-116. Given that the application of a treatment effect parameter to a baseline curve will always be more restrictive than fitting independent curves to each treatment group, please explain why the decision to adopt or not adopt a joint model (including a treatment covariate) was not made *a priori*. Also, please explain why the decision to adopt or not adopt an explicit cure-based model was not made *a priori*.

B15. Company submissionS, section B.3.3.1, page 100. Does the "floor" for the RFS hazard use the general population mortality, or is this weighted by the 4-fold increase? If it is weighted, please comment on the appropriateness of this approach given that a proportion of patients do not receive HSCT?

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

B16. Company submission, section B.3.3.1, Tables 37, 40, 41, 42, 43, and 44. Please explain what is meant by “moderate”, “good”, and “poor” for all columns in which these subjective judgements appear. For example what does “moderate” treatment effect mean? In addition, please explain how the grading of each of these columns affected your choice of parametric survival model and clarify who made these judgements.

B17. Company submission, section B.3.3, figure 31(page 104). . Please provide reasons why curves have been excluded from consideration within this figure and explain why the data shown do not match those presented in Figure 30.

B18. Company submission, section B.3.3. page 103. Please clarify why OS curves which crossed RFS were excluded from further consideration when an alternative explanation of why the curves were crossing was that people were not relapsing in this period of the extrapolation? Please also clarify why, given this argument, the model still applies a logical consistency constraint which minimises RFS when the hazard exceeds that of the OS survivor function. Please also clarify the logic of this model selection criterion given that the base case PSA includes a proportion of samples where the RFS and OS curves cross.

HRQoL

B19. PRIORITY. Company submission, section B.3.4 (page 116). Please provide further details regarding the GLM/GEE model:

a. Were other statistical model forms considered?
b. Was there any control for clustering?
c. What was the distribution family for the data and link function of the GLM/GEE?
d. Please clarify why HSCT status was not included as a covariate in the GLM/GEE?

B20. PRIORITY. Company submission, section B.3.4 (page 116). Please provide further details regarding the post-relapse EQ-5D estimates from BLAST. This should be presented as a table which includes the number of observations and the mean utility post-relapse at each timepoint.

B21. Company submission, section B.3.4, Table 47 (page 120). Please clarify why the age-adjusted utility formula published by Ara and Brazier (Value in Health, 13(5), Figure 2) was not used.

Costs

B22. PRIORITY. Company submission, section B.3.5.4, page 128. Please clarify with details why the online survey responses demonstrated that some of the evidence survey respondents did not understand the exercise. Please clarify why only 2 experts were used to estimate health care resource use.

B23. Company submission, section B.3.5, Table 49, page 122-123. Please provide the source for the cost of a pump applied in the model.

B24. Company submission, section B.3.5, page 122. Please clarify whether the economic analysis accounts for the days that a pump was not allocated to a person receiving blinatumomab within its 5-year lifespan.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

B25. Company submission, section B.3.2.2, page 89. The model assumes there are no disease-related costs after 5 years. Please justify this assumption

B26. Company submission, section B.3.5.4, page 126. Please provide the filename in the reference pack, a PDF or clear web link to reference 105 of the CS.

B27. Company submission, section B.3.5.4, page 127. Given the differences between the populations under appraisal, please comment on the appropriateness of taking the costs of salvage therapy from the previous STA of blinotumomab for people with previously treated B-precursor acute lymphoblastic leukaemia.

B28. Company submission, section B.3.5.4, page 127. Please provide a source for the 37% of people who receive first-line salvage therapy who go on to receive second-line salvage therapy. What was the time to second-line salvage therapy (for those who received this treatment) used in this model.

B29. Company submission, section B.3.5.4, page 127. Please comment on the appropriateness of assuming that the same chemotherapy regimen is given upon a second relapse in the population who receive salvage therapy?

Model

B30. PRIORITY. Company submission, section B.3.2, page 86. Given the need to track HSCT in order to estimate both costs and health outcomes, please clarify why the model has been implemented as a partitioned survival model rather than a state transition model.

B31. PRIORITY. Model. Please explain how to use the model to estimate the total number of HSCTs pre-relapse and post-relapse over the lifetime of the model cohort.

B32. PRIORITY. Model, “Blin Calc” worksheet, GQ9:HW129 & Model, “SOC Calc” worksheet, GQ9:HW129 Please clarify the logic of all calculations used to approximate HSCT receipt and it’s associated treatment over time.

B33. PRIORITY. Model, “Cost Inputs” worksheet, F109:F110. Please clarify the time period for which the probability of receiving HSCT upon relapse (stratified by prior HSCT status) was calculated.

B34. Company submission, section B.3.6, Table 55, page 131. Please clarify why the PSA does not include any uncertainty around MRD response for the SoC group. Please also clarify why the PSA does not include any uncertainty around the proportion of RFS events that are deaths.

B35. Company submission, section B.2.6.1, page 59 and company submission, section B.2.9.2, Table 27, pages 67-68. Given the importance attributed to the 100-day mortality associated with HSCT in the clinical section of the CS, please clarify why this effect was not explicitly included in the health economic model.

B36. Company submission, section B.3.5.4, page 126. Please clarify whether the 38.4% of people in the SoC arm receiving HSCT after 4 years refers only to the post-matching population.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

B37. Company submission, section 3.5.4, page 126. Why were data on post-relapse HSCT not available in either BLAST or the historical control study?

B38. Company submission, section 3.5.4, page 126. If you haven't done so in response to a previous question, please clarify how the probabilities that a patient received HSCT upon relapse were calculated from the data in BLAST and Clinicaltrials.gov: NCT02003612.

B39. Company submission, page 127. Please clarify how the £16,175 per line of salvage chemotherapy was calculated. In the publicly available CS associated with the referenced FAD (Amgen Ltd, Blinatumomab for previously treated B precursor acute lymphoblastic leukaemia: Company evidence submission, page 164, Table 5-15) only a cost of £13,438 per cycle of FLAG-IDA was available.

B40. Model, "Blin Calc" worksheet, cells EQ9:EQ12. Please clarify why the daily pro-rated pump cost and the annual maintenance cost of the pump divided by 365 are applied to the number of outpatient treatment days (cells EC9:EC12) rather than the number of days within a treatment cycle.

B41. Model, "PSA Bootstrap Inputs" worksheet & Model, "PSA Results" worksheet, cell E4. Please clarify why fewer bootstrap samples of the parametric distribution parameters are included compared with the PSA samples (1,000 bootstrap samples versus 10,000 PSA samples).

B42. Model, "Blino calcs" worksheet, columns AM: AN. The (uplifted) general probability of death is greater than 1.0 for later ages. Please comment.

B43. Model, "PSA Inputs" worksheet, I104:I122. Please clarify how the standard errors for the calibrated 6 monthly probabilities of receiving HSCT pre-relapse were calculated.

B44. Model, "Blino calcs" worksheet, columns AW and BA. The mortality probability does not change at the integer age. Please clarify.

EoL

B45. Company submission, section B.2.9.4, page 75 and Company submission, section B.2.13.3, Table 32, page 85. Please explain why the OS figures are quoted based on the matched population of the historical control rather than the unweighted data from this source?

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

Single Technology Appraisal (STA)

Blinatumomab for acute lymphoblastic leukaemia for people with minimal residual disease activity in remission [ID1036]

Dear Gavin and Kawitha,

The Evidence Review Group, ScHARR, and the technical team at NICE have looked at the submission received on 31 October 2017 from Amgen. In general they felt that it is well presented and clear. However, the ERG and the NICE technical team would like further clarification on the clinical and cost effectiveness data (see questions listed at end of letter).

The ERG and the technical team at NICE will be addressing these issues in their reports.

Please provide your written response to the clarification questions by **5pm** on **7 December 2017**. Your response and any supporting documents should be uploaded to NICE Docs: https://appraisals.nice.org.uk/request/38852

Two versions of this written response should be submitted; one with academic/commercial in confidence information clearly marked and one from which this information is removed.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, and all information submitted under 'academic in confidence' in yellow.

If you present data that is not already referenced in the main body of your submission and that data is seen to be academic/commercial in confidence information, please complete the attached checklist for in confidence information.

Please do not 'embed' documents (i.e. PDFs, spreadsheets) within your response as this may result in your information being displaced or unreadable.

If you have any queries on the technical issues raised in this letter, please contact Sana Khan, Technical Lead (Sana.Khan@nice.org.uk). Any procedural questions should be addressed to Thomas Feist, Project Manager TACommA@nice.org.uk.

Yours sincerely

Eleanor Donegan
Technical Advisor – Appraisals
Centre for Health Technology Evaluation

Encl. checklist for in confidence information

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

Dear Eleanor,

We wanted to take the opportunity to thank the NICE Technical Team and ScHARR for their review of our submission and the clarification questions asked. We have endeavoured to answer all queries as fully as possible in the timescale permitted and have provided our responses below.

In addition to this document, we have also provided supporting documentation (word documents, excel files, model code) which I have summarised below.

- **QA6** – a separate word document has been provided with an updated table of all excluded articles in the systematic literature review
- **QB4** – the code used to generate the survival distributions has been provided in a separate file. A separate excel workbook has also been provided with formulas used for all distributions and detailed descriptions regarding the labelling of descriptions.
- **QB8** – the requested sensitivity analyses have been provided in a separate excel workbook
- **QB26 – t**he required reference has been provided as a PDF

Further to this, we have provided a copy of our original submitted model with inclusion of the scenario analyses requested in **QB8** as well as an updated version where have implemented some changes in response to specific clarification questions:

- Mortality capped at 100% (**QB35**)
- Pump cost included for all days after first inpatient stay (**QB24**)
- Utilities applied from Ara & Brazier paper (**QB21**)
- Post-relapse allo-SCT not initiated after 5 years (**QB33**)

We are happy to address any further queries that may arise.

Kind regards

Gavin

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

Section A: Clarification on effectiveness data

Scope/General

A1. PRIORITY. Please provide the proposed wording of the marketing authorisation for blinatumomab for the indication under appraisal.

In alignment with the draft SPC (provided), the anticipated wording for the marketing authorisation is as follows:

BLINCYTO is indicated for the treatment of adults with minimal residual disease (MRD) positive B precursor ALL.

A2. PRIORITY. Company submission (CS), Table 1 decision problem, page 15. The CS states "*Although the cost effectiveness evidence does not consider the Ph+ population or later remission states, due to the substantial unmet need across all sub-populations, blinatumomab should be considered for use in alignment with its full anticipated marketing authorisation.*"

- Please comment on the expected clinical effectiveness and cost-effectiveness of blinatumomab for patients in second or subsequent haematological CR.

As stated in response to Question **A1**, the anticipated license for blinatumomab for the treatment of MRD in patients with positive B precursor ALL is inclusive of latter remission states (ie. CR2 +). In our submission dossier, we present cost-effectiveness results only in the CR1 population (ie. patients in first complete remission), primarily for the following reasons:

- There was limited available evidence in this rare disease to inform estimates of comparative efficacy and cost-effectiveness in latter remission states (please see response to **A14** for more information). When looking within the subset of blinatumomab-treated subjects in CR2 in BLAST, MRD responders did experience benefit from blinatumomab compared to non-responders (Table 1). However, it is also clear from the clinical evidence available in BLAST (presented in our submission dossier) that subjects in CR1 had a better outcome than subjects in CR2; unfortunately, in the absence of robust comparative efficacy data in this small population it has not been possible to evaluate the cost-effectiveness of blinatumomab specifically in CR2+ patients.

Table 1: Summary of OS and RFS for blinatumomab-treated patients in CR2 (BLAST)

CR2 Subpopulation, BLAST	MRD Responders	MRD Non-Responders
RFS, median (months)	[redacted]	[redacted]
OS, median (months)	[redacted]	[redacted]

- Feedback from UK clinical experts indicate that blinatumomab would be used as early as possible in the treatment pathway, at first remission, thus the CR1 population represents the most appropriate ICER for decision making. The recent approval of blinatumomab for the treatment of relapsed/refractory B-cell Ph- ALL (NICE TA450) means that blinatumomab is available as an option for patients who relapse and would likely be used as early as possible at first salvage. We stated in our submission that blinatumomab in the MRD setting should be considered for use in alignment with its full anticipated marketing authorisation as this small population in later remission states *currently* has a

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

significant unmet need; however, over time these patients should effectively be managed with early blinatumomab use in the relapsed/refractory setting. Therefore, it is for patients who are MRD+ in CR1 for whom there are no established treatment options available in current clinical practice.

Indeed, in our submission we present a key scenario analysis whereby use of blinatumomab as first salvage was included for patients who relapse on SoC chemotherapy (Table 2). We believe this analysis would better reflect NICE guidance TA450 and UK clinical practice. As such we consider this to be an alternative base case to the primary base case analysis which does not include the use of blinatumomab for patients who relapse on SoC chemotherapy.

Table 2. Overview of key results from economic evaluation

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	Incremental ICER (£/QALY)
Base Case Analysis					
SoC	[redacted]	[redacted]			
Blinatumomab	[redacted]	[redacted]	**84,259**	**2.95**	**28,524**
Alternative Base Case* – blinatumomab as salvage tx for SoC					
SoC	[redacted]	[redacted]			
Blinatumomab	[redacted]	[redacted]	**33,473**	**1.91**	**17,420**

Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years

These results suggest that the use of blinatumomab in patients with MRD+ B-precursor ALL is a cost-effective use of healthcare resources with an ICER <£30k per QALY; furthermore, when considering use of blinatumomab as a first salvage therapy – aligned with NICE TA450 and expected clinical practice – the ICER decreases below the £20k per QALY threshold. These results are consistent with clinical opinion that treatment with blinatumomab should occur as early as possible in the treatment pathway.

**Note: We have only presented the deterministic ICER for this key scenario analysis in our submission. Additional results can be provided if required (eg. sensitivity analyses)*

- Please comment on the expected clinical effectiveness and cost-effectiveness of blinatumomab for patients with Ph+ disease.

Philadelphia chromosome-positive (Ph+) ALL is a genetically, biologically, and clinically distinct subtype of B-cell precursor ALL. Treatment of Ph+ ALL patients who are resistant to or relapse after first-line therapy remains challenging and outcomes remain poor.

In BLAST, patients with Ph+ ALL were excluded from the study if they were eligible for treatment with TKIs (ie, Philadelphia chromosome-positive patients with no documented treatment failure of or intolerance/contraindication to at least 2 TKIs). In Study MT103-202, TKIs registered for treatment of bcr-abl–positive B-lineage ALL were permitted as concomitant treatment. The numbers of Philadelphia chromosome-positive subjects in both studies was small: a total of 10 subjects (MT103-202, n = 5; BLAST, n = 5).

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

Sub-population analyses for the combined data from Studies MT103-202 and BLAST indicated that the majority of subjects had an MRD complete response regardless of Ph status: 87% (107 of 123) of Philadelphia chromosome negative subjects and 70.0% (7 of 10) of Philadelphia chromosome-positive subjects, who are the most difficult to treat subjects with poor prognoses. Median RFS and duration of hematologic remission from Studies MT103-202 and BLAST was shorter in subjects who were Philadelphia chromosome-positive; however, 95% CIs were overlapping even though the number of Philadelphia chromosome-positive subjects included in the studies was small.

As discussed in response to **A14**, Ph+ ALL patients were not included in the historical control thus there was limited evidence available in this rare subpopulation to inform estimates of comparative efficacy (estimated 15 eligible patients in UK). Although we recognise the challenge of making a recommendation in a population without robust estimates of the cost-effectiveness, this small population has a significant unmet need and has demonstrated comparable response to the larger Ph- population for which the base case analysis addresses.

A3. PRIORITY. Please clarify why the dosing regimen in the licence authorisation is 28µg/day, given that the dosage employed in BLAST was 15µg/m^2.

The rationale for the clinical dose selection for the treatment of MRD+ BCP-ALL is based on the totality of pharmacokinetic (PK), pharmacodynamic (PD), efficacy, and safety information. For the treatment of MRD+ BCP-ALL, the clinical dose tested in Studies MT103-202 (pilot) and MT103-203 (BLAST) was 15 µg/m^2/day for 4 weeks followed by a 2-week treatment-free period between cycles. This regimen was found to be safe and effective in these two studies. Although an equivalent fixed dose regimen of 28 µg/day was not directly tested in these clinical trials, similar exposure levels were expected and had been demonstrated with either the body surface area (BSA)-based dosing or fixed dosing at an equivalent dose in other studies regardless of indications. This is supported by observed steady state concentration (C_{ss}) data, in which the PK was evaluated over a BSA-based dose range from 5 to 90 µg/m^2/day (Studies MT103-104, MT103-202, MT103-203, and MT103-206) and over a fixed dose range from 9 to 112 µg/day (approximately equivalent to 5 to 60 µg/m^2/day; Studies MT103-211, 00103311, 20120216, and MT103-208).

The recommended dose for the treatment of MRD-positive ALL is a cIV infusion at 28 µg/day for 4 weeks followed by a 2-week treatment free period between cycles. This fixed dose regimen is anticipated to be easier to implement in clinical practice compared to the equivalent BSA-based dosing regimen.

A4. Please provide a PDF file of the following reference (reference 5 in CS)
Gökbuget N, Dombret, H., Bonifacio, M., Reichle, A., Graux, C., Havelange, V., Buss, E. C., Faul, C., Bruggemann, M., Ganser, A., Stieglmaier, J., Wessels, H., Haddad, V., Zugmaier, G., Nagorsen, D., & Bargou, R. C. BLAST: A Confirmatory, Single-Arm, Phase 2 Study of Blinatumomab, a Bispecific T-Cell Engager (BiTE®) Antibody Construct, in Patients with Minimal Residual Disease B-Precursor Acute Lymphoblastic Leukemia (ALL). Blood 2014;124:379.

This reference refers to an oral abstract presented at the American Society of Hematology's 56th Annual Meeting, held December 6–9th 2014. A PDF is currently unavailable, but the abstract is accessible online at the following address:

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

http://www.bloodjournal.org/content/124/21/379/tab-article-info

Literature searching

A5. Company submission, Appendix D Identification, selection and synthesis of clinical evidence:

- The searches were last conducted on 19th May 2017. Please update the search and confirm that no relevant studies have been found since then.

We acknowledge that it is best practice to present the most contemporary searches to ensure that all relevant evidence is captured, however we are confident that no further studies to inform this submission have been published in the interim. Moreover, the timeframe for the conduct of our searches (less than 6 months prior to submission) is consistent with previous submissions.

- Please state the exact dates and terms used to search for conference proceedings.

The searches were conducted on June 7th and 8th of 2017.

The American Society of Hematology (ASH), European Hematology Association (EHA), American Society of Clinical Oncology (ASCO), and 2016 American Society for Blood and Marrow Transplantation (ASBMT) conference proceedings are indexed in Embase; thus, the same search terms implemented for the database search in Embase were used. For The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) and The European Cancer Organisation (ECCO)/European Society for Medical Oncology (ESMO), search terms for “acute lymphoblastic leukemia” or “acute lymphocytic leukemia” were used, and all resulting hits were reviewed. The website for the 2017 ASBMT conference was not easily searchable, so search terms for “MRD” or “minimal residual disease” were used in this case.

- Table 73, page 186: Other limits/considerations used in the study selection process across the clinical efficacy/safety SLR suggests that registries were searched. If trial registries were searched, please state the source including search strategies. Please give reasons if trial registries were not searched.

We searched the following clinical trial registries, which are considered to be the most extensively indexed:

- ClinicalTrials.gov (http://www.clinicaltrials.gov/)
- World Health Organization International Clinical Trials Registry Platform (http://www.who.int/ictrp/en/)

The registries were searches using the following terms: “acute lymphoblastic leukemia” or “acute lymphocytic leukemia” and searches were conducted on June 7th and 8th 2017.

- Given that no adverse event studies were identified for blinatumomab for treatment of acute lymphoblastic leukaemia (ALL) with minimal residual disease from the searches presented in Appendix C, please explain why separate adverse events searches were not carried out for blinatumomab only (for information, searches for blinatumomab gives 233 records in PubMed alone).

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

Currently, there is still debate concerning the capacity of clinical studies (randomised controlled trials and observational studies) to yield reliable quantitative estimate of adverse reactions. However, adverse events were reported in the BLAST and are included in Section B.2.10 of the submitted dossier. As is consistent with best practice methods,[2] the primary source of adverse events for HTA assessors is regulatory authorities' documentation (EPAR).[3]

As such, we did not conduct a search systematic search specifically for adverse events.

Systematic review process

A6. Company submission, Appendix D, page 183. Please clarify why three articles were excluded (Figure 46, page 188) for not evaluating treatment of interest, when the inclusion criteria (Table 72, page 185) does not exclude any interventions. Also, given that there is so little evidence available for blinatumomab, please explain why studies with less than 10 patients were excluded.

We acknowledge that the 3 studies excluded due to treatment was an error, and these studies have been re-assessed for inclusion in the review. All 3 studies were excluded for including an irrelevant population.

Please see the accompanying document included in our response for the updated table of excluded studies.

We applied a limit of 10 patients in order to identify the most robust studies to inform decision making. This is consistent with our previous submissions and was intended to exclude 'case-series' which are especially vulnerable to selection bias. However, we have re-appraised all studies excluded by this limit and found no additional relevant studies.

A7. Company submission, Appendix D, page 183. Please confirm how many reviewers conducted data extraction and quality assessment?

To clarify, data extraction and quality assessment was performed according to NICE requirements; two reviewers independently extracted the data and performed the quality assessment. Disputes were resolved by a third and more senior reviewer.

Comparative effectiveness

A8 PRIORITY. Company submission, section B.2.9.4, page 70. With respect to the inverse probability of treatment weighting (IPTW) approach used for propensity score adjustment:

- Please justify why this method has been used over other possible methods that use observational data to inform treatment effectiveness.

Propensity score methods were applied to control for key prognostic factors. The propensity score approach can potentially create a balance between the blinatumomab-treated patients in the MT103-203 study and the historical comparator patients in the 20120148 study with respect to multiple clinical factors that are thought to affect a patient's general prognosis. Such a balance, if adequately achieved, would allow for valid statistical inferences.[4] More recently, they have been used in a regulatory setting when needing to evaluate results from non-randomized studies for regulatory decision making.[5, 6] The inverse probability of treatment weighting (IPTW)

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

approach for propensity score adjustment was chosen to estimate treatment effectiveness because, given time-to-event endpoints, alternative methods such as stratification or covariate adjustment have been shown in the literature to produce biased estimates of marginal and conditional hazard ratios.[7, 8]

- Please clarify how the assumption of selection on observables was assessed.

One of the main assumptions for propensity score analysis is that the assignment of the treatment is independent of the outcome conditional on the covariates. This requires that all relevant confounders are observed and included as candidate variables for the propensity score model.

For this analysis, candidate variables were selected based on lengthy discussion among study team experts and clinicians; the majority of the covariates were chosen based on prognostic factors that have been identified for ALL in published literature and to account for potential regional differences in treatment practices. Candidate variables are those that are common to both the databases and are thought to be important for characterising the blinatumomab-treated population. We acknowledge the limitation that, unlike with randomised studies, propensity score analysis does not tend to create a balance with respect to all covariates (including unmeasured and unknown covariates). However, we feel that we have adequately populated the covariate set with all available relevant prognostic factors to provide a valid statistical comparison between blinatumomab and control populations.

- Please clarify how the assumption of overlap between the two studies was assessed. Is this considered to be justified for average treatment effect (ATE) and average treatment effect on the treated (ATT) weights?

The assumption of overlap is equivalent to the condition that the probability of receiving treatment is non-zero for all subjects. This assumption is required to obtain accurate estimates for the unobserved counterfactual means in ATE and ATT.

We assessed the overlap and balance between the two populations using a variety of methods. Upon deriving propensity scores (PS) for each patient, balance between the two treatment groups with respect to their PS were assessed via box plots. The overall balance was considered to be sufficient given that at least 25% of the historical data overlapped with the inner 95th percentile of the blinatumomab data, as pre-specified.

With respect to individual covariates considered for the propensity score model, two methods were employed to ascertain the balance between the data sources before and after propensity score adjustments. The first method involved univariate regression models with the baseline factor as the dependent variable and the treatment group as the independent variable. For categorical factors, a logistic regression model with robust variance estimation was used. For continuous variables, a general linear model with robust variance estimation was used. The p-value associated with the treatment group effect from each model was used to compare the before- and after-effects of the PS adjustment.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

The second method involved calculation of standardised differences. Standardised differences can be used to ascertain the balance in a way that is not dependent on the sample size. Criteria for deciding whether the balance was adequate included: univariate p-values greater after adjustment and not considered significant and a standardised difference less than at least 0.20 (best balance achieved when less than 0.10).

If important covariates or baseline factors were not adequately balanced upon doing the evaluations described above, and the covariate was considered prognostic with respect to the endpoint, then those factors may be added as additional covariates to the endpoint analysis model for sensitivity analyses.

- A subgroup of BLAST (n=73), was used to estimate comparative effectiveness. Please clarify why each of the listed criteria was required (as opposed to using the whole BLAST trial population).

Since the historical comparator study included only a very limited number of patients beyond their first remission (>CR1), and since remission duration decreases significantly with an increasing number of prior relapses, the primary analysis set for the propensity score analysis was defined such that only those in CR1 were included. This is the primary criterion that reduced the size of the blinatumomab group from the full BLAST population (N=116) to that for the primary analysis set for the propensity score analysis (N=73). The full set of criteria for the Primary Analysis Set are shown below:

Study MT103-203 criteria:

- Received any infusion of the investigational drug.
- Philadelphia negative B–precursor ALL in complete haematological remission defined as less than 5% blasts in bone marrow after at least three intensive chemotherapy blocks.
- MRD-positive at a level of $\geq 1 \times 10^{-3}$ (PCR only in Study MT103-203) but otherwise in complete haematological remission
- At least 18 years old at the MRD baseline date
- In their first remission (CR1 only)

20120148 criteria:

- Philadelphia-negative B-precursor ALL in complete haematological remission
- MRD-positive at a level of $\geq 1 \times 10^{-3}$ regardless detection method
- At least 18 years old at the MRD baseline date
- Time to relapse greater than 14 days from date of MRD detection

- Please clarify whether the assumption of ignorability of treatment is considered to be justified for ATE and ATT weights.

The ignorability assumption states that the assignment of the treatment is independent of the outcome conditional on the covariates. This assumption is also referred to as the selection on observables (see part a).[9] ATE and ATT are only identifiable if sources of selection bias, such as violating the assumption of ignorability, are eliminated. If all relevant covariates are considered for the propensity score model, and the resulting model provides adequate balance between the

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

populations, then ATE and ATT counterfactuals are estimable and the assumption of ignorability holds. For this analysis, we accounted for all relevant prognostic factors and achieved adequate balance, therefore we conclude that the ignorability assumption holds, and ATE and ATT are estimable.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

A8 Additional Query. CS Table 29 page 73. The baseline characteristics of the SOC and Blinatumomab groups are presented before and after propensity score adjustment with ATT weights. Could you please confirm that the figures presented in the table, and specifically the sample size (N=174.3 control, N=78.5 blinatumomab) are correct? The formulae presented in Appendix L page 219 indicate that the under the ATT assumption the treatment arm should not be weighted (weight=1). The curves presented in the cost effectiveness section, Figure 22, page 97, indicate N=155.1 control, N=73 blinatumomab. This relates to question **A8** in the original clarification letter.

There are two issues that we would like to address/correct in this response: **1)** Reporting of sample size for stabalised IPTW analysis; **2)** Application of weights in cost-effectiveness analysis.

Reporting of sample size in stabilised IPTW analysis

The sample size presented in Table 29 of the submission were unfortunately reported incorrectly – please find updated Table below with the correct sample size reported. It is important to note that these values reflect the baseline characteristics after the propensity score adjustment using the *stablised* IPTW (sIPTW) and as a result the weighting for the blinatumomab arm is not equal to one.

Characteristic	Unweighted				Stabilised IPTW			
Mean (SD)/n (%)	Control [redacted]	Blinatumomab [redacted]	Standard Difference	P-value	Control [redacted]	Blinatumomab [redacted]	Standard Difference	P-value
Age at primary diagnosis (years)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gender (Female)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Country (Not Germany)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
MRD at Baseline (recoded)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Time from diagnosis to	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

baseline (months)								
WBC at diagnosis (>30,000/mm^3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
WBC at diagnosis (continuous, log10)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
T411mll4 mutation (Yes)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Prior chemotherapy (GMALL)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Stabilised IPTW can be conducted to reduce the potential instability caused by very large weights' and was applied for this propensity score analysis. In order to calculate stabilised weights, the IPTW is multiplied by the marginal probability of receiving the actual treatment received (Cole and Hernan, 2004).

$$sw_{j,k} = w_i \frac{n_k}{\sum_{k=1}^{2} n_k}$$

Where $sw_{j,k}$ represents the sIPTW for the jth subject from treatment k and n_k represents the sample size for treatment k. This results in non-integer values for both blinatumomab and SoC in this analysis.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

Application of weights in Cost-effectiveness analysis

In the cost-effectiveness analysis, the ATT propensity score analysis was applied using non-stabilised ATT weights. Unfortunately, this was done in error and lacks consistency with the presentation of clinical results.

However, we are confident that the application of stabilized ATT weights would have no impact on the analyses conducted for cost-effectiveness since the relevant size of the two groups are the same (73:151=0.47; 20.9:44.4=0.47). We would be happy to provide an updated version of the calculations to confirm consistency of the results but have been unable to complete this in the timeframe – however, this can be provided if required.

A9 PRIORITY. Company submission, section B.2.9.4, page 71. Candidate variables for the propensity score model are provided. Please also provide the final used model.

The final model has been summarised in the table below.

Table 3. Summary of Propensity Score Model Covariates (Primary Analysis Set)

	Primary Analysis Set (N=255)		
	Estimate (SE)	Wald Chi-Square Statistic	p-value[a]
Age at primary diagnosis (years)	[redacted]	[redacted]	[redacted]
Time from diagnosis to baseline (months)	[redacted]	[redacted]	[redacted]
MRD level at baseline	[redacted]	[redacted]	[redacted]
Type of prior chemotherapy			[redacted]
Not GMALL	[redacted]	[redacted]	
Time from diagnosis to baseline (months) x Type of prior chemotherapy			[redacted]
Not GMALL	[redacted]	[redacted]	

Footnotes: [a]p-value represents the statistical significance of including the covariate (or interaction term) into the model.
Abbreviations: N = Number of subjects in the analysis set

A10. Company submission, figures 17 and 18, pages 74 and 75. The included figures indicate time from randomisation to event. Please clarify whether the figures and related statistics relate to time from first blinatumomab treatment for BLAST patients and 14 days after the MRD baseline date for patients from the historical comparator study 20120148.

We apologise for the mislabelling of the x-axis – to clarify, the baseline date (i.e. start time) for relapse-free and overall survival presented in the submission was defined as 14 days after the MRD baseline date for the historical comparator study (20120148) and the date of the first blinatumomab treatment for BLAST patients.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

A11 PRIORITY. Company submission, appendix L, pages 216. The start date for time to event outcomes was defined as "*14 days after the MRD baseline date for 20120148 patients and the date of the first blinatumomab treatment for BLAST patients*"

- Please justify why the median time from MRD detection to first blinatumomab dose is the most appropriate cut-point to use

To correctly compare RFS and OS, a baseline date must be well aligned between the two populations. Aligning both populations using their MRD detection date would lead to an immortality bias for MT103-203 patients due to the fact that patients relapsing or dying after MRD detection would not have been included in the MT103-203 study or treated with blinatumomab. To remove this bias, initially two different baseline dates were used: the date of first blinatumomab treatment for MT103-203 patients, and the date of MRD detection for 20120148 patients. However, additional bias is introduced in that 20120148 patients with rapid relapse following MRD detection would not have counterparts in the MT103-203 study. To better align the populations and reduce bias due to the definition of MRD baseline date, study 20120148 patients were excluded if their time to relapse was less than 14 days, which is the median time between MRD detection and first blinatumomab dose for MT103-203 patients, and the baseline date for the 20120148 control population was set at MRD detection date plus 14 days.

The start time for relapse-free and overall survival was defined as 14 days after the MRD baseline date for 20120148 patients and the date of the first blinatumomab treatment for MT103-203 patients. For study 20120148, MRD baseline date was defined as 14 days after the MRD detection date following complete remission after at least three blocks of chemotherapy: the time point of eligibility had these historical patients been screened for study MT103-203. Because study 20120148 captures an extended disease history, some patients might have multiple MRD detection dates following multiple complete remissions from multiple lines of chemotherapy; at each of these dates the patient would have been eligible for study MT103-203 (provided they had at least three total blocks of chemotherapy). For these patients, the MRD baseline date was defined as 14 days after the date of first MRD detection following the first complete remission.

- Please provide further statistics for this variable e.g. what was the range of values observed for the time between MRD detection and first blinatumomab dose? How many RFS/OS events occurred during the first 14 days, leading to exclusion from the historical control data set?

A summary of the descriptive statistics for time from first documented MRD-positive test to the 1st dose of blinatumomab is provided below.

Table 4. Descriptive Statistics for Time from 1st Documented MRD-positive Test to the 1st Dose of Blincyto (Full Analysis Set): Study MT 103-203

	Time from 1st documented MRD-positive to 1st dose of blinatumomab (Full Analysis Set)* (N=116)
n	[redacted]
Mean (days)	[redacted]
Median (days)	[redacted]

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

	Time from 1st documented MRD-positive to 1st dose of blinatumomab (Full Analysis Set)* (N=116)
Q1 (days)	[redacted]
Q3 (days)	[redacted]
(Min, Max) (days)	[redacted]

***Time from 1st MRD to first blinatumomab dose is calculated as First MRD date - First blinatumomab dose date +1.**

Abbreviations: N=Number of subjects in the analysis set.
Footnotes: n is the number of subjects with MRD positive results (≥ 0.001) at the central lab at the baseline.

A total of [redacted] patients were excluded from the historical control data set due to relapse during the 14 days after MRD baseline.

A12. Please provide details of prior ALL treatments received by patients in the BLAST study.

Available details on the prior ALL treatments received by patients in BLAST is summarised below.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

Table 5. Previous Anti-Tumour Drug Treatment

Characteristic Category	Full Analysis Set (N=116)
Maximum line of therapy	
Front line treatment	[redacted]
First relapse treatment	[redacted]
Second relapse treatment	[redacted]
Front line treatment	[redacted]
Pre-phase	[redacted]
GMALL	[redacted]
combination of regimen /other	[redacted]
GMALL elderly	[redacted]
GRAALL	[redacted]
UKALL	[redacted]
GIMEMA	[redacted]
PETHEMA	[redacted]
FLAG-Ida	[redacted]
NILG	[redacted]
TKI	[redacted]
FRAALLE	[redacted]
Hyper-CVAD	[redacted]
iBFM	[redacted]
AIEOP	[redacted]
HOVON	[redacted]
ALL-2009	[redacted]
ALL-2009 elderly	[redacted]
EWALL elderly	[redacted]
GRAAPH	[redacted]
LALA94	[redacted]
Romanian Group for ALL	[redacted]

Abbreviations: N=Number of subjects in the analysis set.
Footnotes: Subjects who received a regimen in combination with other study group specific treatments are counted in each category.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

Table 6. Previous Anti-Tumour Drug Treatment

Characteristic Category	Full Analysis Set (N=116)
First relapse treatment	[redacted]
Other relapse regimen	[redacted]
FLAG-Ida	[redacted]
Hyper-CVAD	[redacted]
TKI	[redacted]
Second relapse treatment	[redacted]
Other relapse regimen	[redacted]
TKI	[redacted]
Number of intensive treatment blocks per patient	
2 blocks	[redacted]
3 blocks	[redacted]
4 blocks	[redacted]
5 blocks	[redacted]
6 blocks	[redacted]
7 blocks	[redacted]
8 blocks	[redacted]
9 blocks	[redacted]
10 blocks	[redacted]
11 blocks	[redacted]
12 blocks	[redacted]
13 blocks	[redacted]
14 blocks	[redacted]
16 blocks	[redacted]
22 blocks	[redacted]
29 blocks	[redacted]

Abbreviations: N=Number of subjects in the analysis set.
Footnotes: Subjects who received a regimen in combination with other study group specific treatments are counted in each category.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

Table 7. Previous Anti-Tumour Drug Treatment

Characteristic Category	Full Analysis Set (N=116)
Front line treatment	
1 block	[redacted]
2 blocks	[redacted]
3 blocks	[redacted]
4 blocks	[redacted]
5 blocks	[redacted]
6 blocks	[redacted]
7 blocks	[redacted]
8 blocks	[redacted]
9 blocks	[redacted]
10 blocks	[redacted]
11 blocks	[redacted]
12 blocks	[redacted]
14 blocks	[redacted]
15 blocks	[redacted]
23 blocks	[redacted]
First relapse treatment	
1 block	[redacted]
2 blocks	[redacted]
3 blocks	[redacted]
4 blocks	[redacted]
6 blocks	[redacted]
7 blocks	[redacted]
Second relapse treatment	
1 block	[redacted]

Abbreviations: N=Number of subjects in the analysis set.
Footnotes: Subjects who received a regimen in combination with other study group specific treatments are counted in each category.

A13. Company submission, page 67. Please clarify whether the historical control study included patients treated in years 2000-2017 (as suggested in the CS), or 2000-2014 (as suggested in the Observational Study Report).

Apologies for this error; the correct patient recruitment period for the historical comparator study is 2000–2014, as recorded in the Observational Study Report.

A14. Company submission, section B.2.9.3, page 68. Please clarify why Ph+ patients and CR2+ patients were not included in the historical control study.

The historical control study utilised retrospective data from 8 study groups, and included CR1 or CR2 patients who had received 3 prior intensive blocks of chemotherapy in the full analysis set

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

(FAS). While CR1 or CR2 patients both met inclusion criteria, all but 2 patients were CR1 at the time of baseline MRD. Therefore, when filtering the historical control study patients for patient characteristics to match the BLAST patient characteristics (direct comparison analysis set), only CR1 patients met these criteria.

Ph+ ALL patients were not included in the historical control inclusion criteria given the evolving paradigms in prior treatment algorithms (i.e. introduction of newer generation tyrosine kinase inhibitors [TKI]) when compared to Ph- ALL patients which were relatively similar over the study period.

A15. Company submission, section B.2.9.3, page 68. Please provide details of any peer reviewed published data from the historical control study. If none are available, please provide further details regarding the study design and enrolment procedures.

No peer reviewed published data from the historical control study are currently available (manuscript in development). However, the historical comparator data has been submitted to FDA and EMA and will be published as part of their assessment.

The historical control study was a retrospective non-interventional cohort study of historical treatment and outcome data from MRD-positive patients with Ph- Bcell- precursor ALL who had received standard of care treatment according to national study protocols. The primary research objectives were to estimate RFS and OS for patients with MRDpositive- B-cell precursor ALL. Assessment of MRD response was not included in the study because of the variability in treatment regimens after documentation of MRDpositive- status (e.g., continued chemotherapy, investigational agent, or no intervention) and variable availability of MRD assessments after the qualifying baseline MRD-positive value.

The study population was assembled from patient databases of ALL study groups in Europe that included MRD testing in their protocols. All subjects who were treated at participating study group facilities, diagnosed with ALL in the year 2000 to 2014, and who met the eligibility criteria were included in the study. Subjects ages 15 years or older with Ph- B-cell precursor ALL in haematologic CR (defined < 5% blasts in bone marrow after at least 3 intensive chemotherapy blocks) were included if MRD was detected at a level of $\geq 1 \times 10^{-4}$ by PCR or $\geq 1 \times 10^{-3}$ by flow cytometry at a reference lab; a history of ALL treatment (including response to first therapy, number of prior relapses) was available; and relapse status and disease follow-up after time point of MRD detection was available. Subjects were excluded from the analysis if they had extramedullary disease at the time point of MRD detection, were exposed to blinatumomab within 18 months of MRD detection, or underwent allogeneic HSCT before MRD detection. Subject data were entered into a study-specific electronic case report form to ensure a standardised data collection process across study groups.

A16. Company submission, Appendix D, Figure 47, page 195. Please modify the figure to include the reasons for not enrolling 95 patients out of the total of 211 patients who were screened.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

Of the 95 patient screen failures, the majority were due to either having a MRD level < 1 x 10^{-3} (which is below the inclusion criteria for this study) or having an overt leukemia relapse (no longer having MRD). Please see the table for the full listing of screen failure reasons.

Table 8. Consolidated Screening Failure: Study MT103-203

Screening Failure	
Summary Reason	**Total**
Active Infection	[redacted]
Alternative Therapy	[redacted]
CD19 Negative	[redacted]
CNS Relapse	[redacted]
Consent Withdrawn	[redacted]
Hepatic	[redacted]
MRD < 10E-3	[redacted]
Neurologic Disorder	[redacted]
Overt Relapse	[redacted]
Technical	[redacted]
Grand Total	95

Section B: Clarification on cost-effectiveness data

Literature searching

B1. Company submission, appendix D, page 183. Published cost effectiveness studies. Given that no cost-effectiveness studies were identified for ALL with minimal residual disease from the searches contained in Appendix C, please explain why separate cost-effectiveness searches were not carried out for ALL only i.e. ALL terms combined with an economic evaluations study design filter.

Search terms and study protocol were designed to best characterise the decision problem for this submission. As such it was felt that indirect evidence from a general ALL population was not informative for this particular decision making, *i.e., adult ALL patients in complete haematological CR with MRD.* We are confident that all appropriate evidence was captured as part of the systematic literature review.

B2. Company submission, appendix D, Table 71, page 185: York Centre for Reviews and Dissemination search algorithm – page 183 of appendix D states that "these databases were searched via the EMBASE®, PubMed, Cochrane Library and York Centre for Reviews and Dissemination". Table 71 shows ALL terms combined with costs/economic terms, but this is not consistent across the other relevant databases i.e. EMBASE®, PubMed, Cochrane Library where it is ALL+MRD and not ALL+ costs/economic terms. Please comment on the likelihood that no economic evaluations have been missed as a consequence of this approach.

The review of the York databases was intended to specifically identify HTAs and cost-effectiveness analyses, which is why terms for cost and economic were included with the string. In order to make the strategy as broad as possible these searches were not limited by MRD

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

specific terms however we did not identify any relevant evidence. Despite this we are confident that our broad approach captured all relevant evidence.

B3. Company submission, appendix E, page 198. HRQoL studies. Given that no HRQoL studies were identified for ALL with minimal residual disease from the searches contained in Appendix C, please explain why a separate HRQoL search was not carried out for ALL only i.e. ALL terms combined with HRQoL terms?

Search terms and study protocol were designed to best characterise the decision problem for this submission. As such it was felt that indirect evidence from a general ALL population was not informative for decision making. We are confident that all appropriate evidence was captured as part of the systematic literature review.

Survival modelling

B4. PRIORITY. Company submission, section B.3.3. pages 93-116. Five model types have been applied within the CS (Unrestricted, Restricted, Restricted Cubic Spline, Mixture cure, and Non-Mixture cure). For the general case, and the best fitting model in each category (e.g. Gompertz for unrestricted)

- Please provide mathematical equations for the model. For example, for the Gompertz unrestricted model, please provide the regression equation "*including treatment -group interaction terms for every distributional parameter*". Please provide the code and output showing how these models have been fitted in the relevant software package (i.e. flexsurv in R for the parametric models, STATA for the cure models)

The code used to generate the survival distributions has been provided along with a separate workbook with the formulas used for all the distributions generated by the code. Regarding the labelling of the distributions, please see the detailed description provided separately in the excel workbook.

B5. PRIORITY. Company submission, section B.3.3, pages 93-116. Please clarify why the lognormal mixture cure model (and presumably every other mixture cure model fitted) does not include the expected mortality for the cured fraction (see Lambert, The Stata Journal, 2007, 7(3) i.e. the source of the STATA procedures cited in the CS). Please also comment on the potential bias associated with estimating survival probabilities for cured patients based on a fixed model start age rather than the observed distribution of patient age within the clinical studies.

We did not include expected mortality in the curve fitting process as such mortality is likely to be immaterial in these patient over the follow-up period over which the distributions were estimated. We therefore do not believe that the omission of this mortality from the estimation procedure will bias the analysis in favour of blinatumomab. We also recognise that there is some potential for bias by estimating outcomes for the mean age rather than the distribution of age, although we had no reason to believe that this bias would favour one treatment or the other. To assess this potential bias, we generated results setting the model start age from 18 to 76 years which were pooled using the distribution of CR1 patients in BLAST. The ICER based on the weighted average incremental costs and QALYs was similar to the base case (£29,174 per QALY).

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

B6. PRIORITY. Company submission, section B.3.3. pages 93-116. Please explain the conceptual implications of selecting a cure model for OS but not for RFS. How can patients who are cured in terms of OS still experience relapse according to RFS?

For the base case, we used a restricted Gompertz model for RFS. While this model does not include an parameter representing the cure fraction, it does yield distributions of RFS with a non-zero asymptote and therefore is effectively a "cure" model. For scenario analysis 3, an RCS log-logistic model was used for RFS. While this is not a cure model, the RCS log-logistic distribution was selected to represent RFS in the less favourable scenario, as it had the third best statistical fit; some supportive evidence in favour of proportional odds based on the treatment effect counterfactual plots, and appeared to represent a plausible lower bound on the benefit of treatment with blinatumomab. Note: for OS a non-cure model was also adopted (RCS Weibull).

Conceptually, we would argue that it is probably more appropriate to use cure models for both RFS and OS as in our base case. It should be noted, however, that it is possible that some patients might not be cured for RFS following initial treatment but would be cured subsequently (e.g., by post relapse transplant).

B7. PRIORITY. CS, Section B.3.3. pages 93-116. Please explain whether and how clinical judgement was used to inform the selection of the parametric functions for RFS and OS in the model.

Clinical input pertaining to the RFS and OS survival of patients with MRD+ ALL was a key aspect informing selection of the parametric functions used, in particular relating to the validation of modelled survival projections. Specifically, UK clinicians were asked to comment on expected survival of patients currently observed in clinical practice (at landmark timepoints), the appropriateness of assuming a cure at a specific timepoint, and the proportion of patients that may realise a cure given current SoC. Clinicians were also asked to comment on the magnitude of benefit likely to be derived from obtaining an MRD-negative status.

Feedback consistently suggested that patients alive after 5 years would be considered cured although an earlier timepoint was also realistic – patients who remain relapse-free after 2–3 years from initial treatment were thought to be at a low risk of relapse and could be considered cured. The cure fractions estimated by the base case modelled projections were considered appropriate (and aligned with expectations in clinical practice). It was also noted that the historical comparator used to inform survival estimates for the standard of care arm was highly likely to be generalisable to the current UK treatment protocols given that no significant changes to clinical practice has occurred in the last decade.

As a result of this feedback, it was considered that the parametric functions used in the base case analysis appropriately capture long-term survival and provide clinically valid estimates to inform the economic evaluation.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

B8. PRIORITY. Company submission, section B.3.3. pages 93-116. Please provide complete sensitivity analyses which include all combinations of functional forms for RFS and OS (assuming same curve type for each treatment group). Please ensure that this analysis does not exclude combinations of OS and RFS where the RFS and OS curves cross (in such instances, please constrain RFS to the minimum of RFS and OS). Please report the results of these analyses in the form of a table including incremental QALYs, incremental costs and ICERs for each comparison.

The requested sensitivity analyses have been provided and are included in a separate workbook.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

B9. Company submission, section B.3.3.1, Figure 23 (page 98), figure 30(page 104), appendix P, figure 58 and 64. Please provide model fit statistics (BIC) in a table rather than a line graph.

Table 9. Model Fit Statistics (OS)

Rank	Distribution	BIC
1	RCS Log-Logistic (R)	1169.49718
2	RCS Weibull (R)	1169.98653
3	RCS Lognormal (R)	1171.03738
4	Lognormal Non-Mixture (Cure)	1171.67567
5	Lognormal Mixture (Cure)	1173.18727
6	Gen. Gamma (R)	1173.34913
7	Lognormal (R)	1173.67129
8	Gen. F (R)	1176.19578
9	Lognormal Non-Mixture (Cure + R)	1176.96428
10	Gamma Non-Mixture (Cure)	1177.0575
11	Lognormal Mixture (Cure + R)	1177.8342
12	Lognormal (U)	1179.17301
13	Log-Logistic (R)	1179.88257
14	RCS Log-Logistic (U)	1180.35114
15	RCS Weibull (U)	1180.6083
16	Gompertz (R)	1181.62989
17	RCS Lognormal (U)	1181.93832
18	Lognormal Non-Mixture (Cure + U)	1182.05656
19	Gamma Non-Mixture (Cure + R)	1182.23081
20	Lognormal Mixture (Cure + U)	1182.96908
21	Weibull Non-Mixture (Cure)	1183.0339
22	Gen. Gamma (U)	1183.7722
23	Log-Logistic (U)	1185.32641
24	Gompertz (U)	1187.0162
25	Weibull Mixture (Cure)	1188.20234
26	Weibull Non-Mixture (Cure + R)	1188.55182
27	Gen. F (U)	1190.68764
28	Weibull Non-Mixture (Cure + U)	1192.72218
29	Weibull Mixture (Cure + R)	1193.66084
30	Gamma Mixture (Cure + U)	1194.83721
31	Pw. Exp.	1196.28864
32	Weibull Mixture (Cure + U)	1197.17415
33	Exponential	1197.45691
34	Weibull (R)	1197.72265
35	Weibull (U)	1201.82174
36	Gamma Mixture (Cure)	Failed to Converge
37	Gamma Mixture (Cure + R)	Failed to Converge
38	Gamma Non-Mixture (Cure + U)	Failed to Converge

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

Table 10. Model Fit Statistics (RFS)

Rank	Distribution	BIC
1	Gompertz (R)	1222.060922
2	Gompertz (U)	1225.586802
3	RCS Log-Logistic (R)	1225.662197
4	Lognormal (R)	1227.202426
5	Weibull Non-Mixture (Cure)	1227.298547
6	Lognormal Non-Mixture (Cure)	1227.459555
7	Lognormal Mixture (Cure + R)	1228.875339
8	Log-Logistic (R)	1228.876169
9	RCS Lognormal (R)	1229.011902
10	Lognormal Non-Mixture (Cure + R)	1229.114534
11	Gen. Gamma (R)	1229.285887
12	RCS Weibull (R)	1230.052249
13	Gen. F (R)	1230.616178
14	Weibull Non-Mixture (Cure + R)	1230.720394
15	Gamma Non-Mixture (Cure)	1231.213658
16	Lognormal (U)	1232.716224
17	Gamma Mixture (Cure + R)	1233.307
18	Lognormal Mixture (Cure)	1233.420019
19	Gamma Non-Mixture (Cure + R)	1233.446829
20	Log-Logistic (U)	1234.357799
21	Lognormal Mixture (Cure + U)	1234.391678
22	Weibull Mixture (Cure + R)	1234.439432
23	Lognormal Non-Mixture (Cure + U)	1234.655132
24	Weibull Mixture (Cure)	1235.512
25	Weibull Non-Mixture (Cure + U)	1235.785318
26	RCS Log-Logistic (U)	1236.037173
27	RCS Weibull (U)	1236.606596
28	Gamma Mixture (Cure)	1236.98509
29	Weibull Mixture (Cure + U)	1238.882405
30	RCS Lognormal (U)	1239.740574
31	Gen. Gamma (U)	1240.160742
32	Gamma Mixture (Cure + U)	1244.3427
33	Gamma Non-Mixture (Cure + U)	1244.415065
34	Gen. F (U)	1244.547913
35	Weibull (R)	1257.919498
36	Weibull (U)	1260.686934
37	Pw. Exp.	1265.063717
38	Exponential	1321.743219

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

B10. Company submission, section B.3.3.1, Table 35, page 94. The CS refers to "unrestricted" parametric models which the ERG understands to mean fitting separate models to each treatment group without the inclusion of a treatment effect parameter (a hazard ratio or a constant acceleration factor), or otherwise relaxing this restriction as implemented in the CS through the inclusion of the interaction terms. However, Table 35 includes a tick mark for an HR or AF in the "treatment effect" column for unrestricted models. Please clarify.

Please see explanation of the parametrisation of treatment effects in the different models in the excel workbook provided.

B11. Company submission, section B.3.3.1, page 95. Please clarify why 8% of people receiving standard of care (SoC) are assumed to achieve MRD response, given that the clinicians consulted suggested this was "no greater than 10%".

In the absence of concrete information other than clinical expert opinion, 8% was selected as a reasonable but conservative estimate of the proportion of patients receiving who might achieve MRD response. The sensitivity of model results to this parameter were explored in a scenario analysis.

B12. Company submission, section B.3.3.1, page 95-97. In 4 sentences or less, and using non-technical language, please explain what the analysis using the Berry external data mean.

The data from the Berry study were used to assess the external validity of the RFS and OS distributions used in the model as well as the magnitude of the increase in RFS and OS that would be expected given the effect of blinatumomab on MRD response. Details of this analysis are provided in Section B 3.3.3.1 of the submission. Results of these analyses suggest that the shapes of the survival distributions and the magnitude of the gains in survival that are projected based on parametric distributions used in the model are consistent with those that would be expected given the differences in MRD response for blinatumomab versus SoC and the data on OS and RFS by MRD response from Berry.

B13. Company submission, section B.3.3.1, page 95-97. Please provide a comparison of the matched OS data to the Berry data, in the same way as was presented for the RFS data (see page 97). These data are discussed but not presented in the CS.

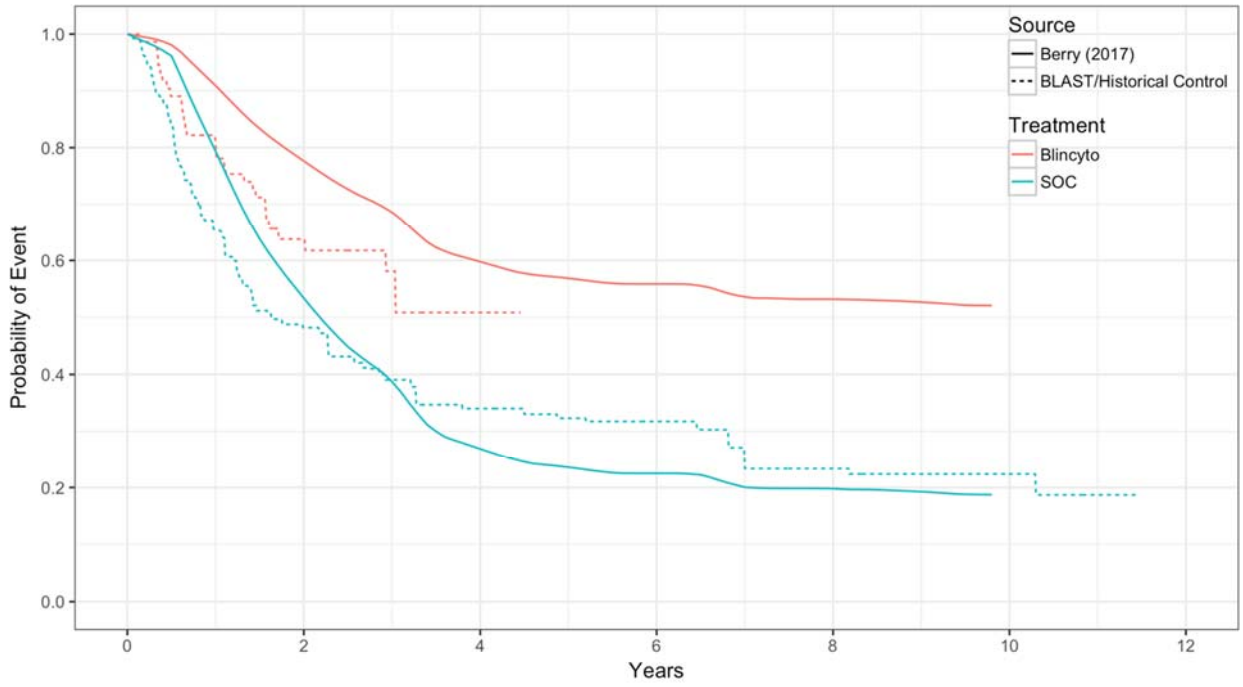
Weighted projections of OS based on Berry et al. are shown alongside OS from BLAST and the historical comparator study in Figure 1.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

Figure 1. Projected OS for blinatumomab and SOC based on OS by MRD response from Berry et al. compared with OS in propensity-matched analysis of BLAST and historical control study



B14. Company submission, section B.3.3. pages 93-116. Given that the application of a treatment effect parameter to a baseline curve will always be more restrictive than fitting independent curves to each treatment group, please explain why the decision to adopt or not adopt a joint model (including a treatment covariate) was not made *a priori*. Also, please explain why the decision to adopt or not adopt an explicit cure-based model was not made *a priori*.

We did not have any priors regarding the appropriate parameterisation of the treatment effects for RFS and OS and were instead guided by the criterion-based selection process, which included consistency with counterfactual treatment effect plots, visual fit, BIC, and consistency with external data. While a model with more parameters – i.e., less restricted – will always fit better than a less restrictive one in the same class, we used BIC as a measure of goodness of fit which penalises models with more parameters and therefore helps avoid the possibility of over-fitting.

Although clinical expert opinion and evidence from the literature suggested that cure models may be the most appropriate way to model this disease area, we did not make a decision to exclude non-cure based models *a priori.* This was primarily due to the fact that we wanted to consider a comprehensive set of models (including cure and non-cure) and to select the best models based on a reasonable set of criteria including our consistency with our priors. Since we ended up using a cure model in the base case analysis (the Gompertz is effectively if not explicitly a cure model), focusing only on cure models *a priori* would not have impacted the distributions we selected.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

B15. Company submission, section B.3.3.1, page 100. Does the "floor" for the RFS hazard use the general population mortality, or is this weighted by the 4-fold increase? If it is weighted, please comment on the appropriateness of this approach given that a proportion of patients do not receive HSCT?

The floor for the RFS (and OS) hazard is the general population mortality adjusted for the 4-fold increase in mortality. Because we lacked data on the increase in mortality for a population of patients with Ph- R/R ALL who may or may not have received HSCT, we used the value for post-transplant patients. It should be noted, however, that our estimate of the increase in mortality was based on an analysis of the long-term consequences of allogeneic HSCT conducted by Martin et al that compared patients who underwent HSCT versus the general population. {Martin, 2010 #11} The precise proportion of excess risk that is due to HSCT versus ALL *per se* is not possible to ascertain. However, a large share of late mortality was due to recurrent malignancy. Furthermore, even those patients who were not exposed to HDT and HSCT would have received at least one course of chemotherapy and potentially radiotherapy. These too might contribute to excess long-term mortality.

B16. Company submission, section B.3.3.1, Tables 37, 40, 41, 42, 43, and 44. Please explain what is meant by "moderate", "good", and "poor" for all columns in which these subjective judgements appear. For example what does "moderate" treatment effect mean? In addition, please explain how the grading of each of these columns affected your choice of parametric survival model and clarify who made these judgements.

As with all visual assessments of goodness of fit, these assessments are based on subjective judgements. Generally speaking, we used the following definitions:

- "Good fit": The two curves are virtually the same, with no systematic over or under estimation
- "Poor fit": The two curves are substantially different with apparent systematic over or underestimation over some range of the curve
- "Moderate": Intermediate between "good fit" and "poor Fit"

Judgement of goodness of fit was initially made by the analyst conducting the regression analyses. These judgements were then confirmed by the team who contributed to the evaluation.

B17. Company submission, section B.3.3, figure 31(page 104). . Please provide reasons why curves have been excluded from consideration within this figure and explain why the data shown do not match those presented in Figure 30.

Figure 31 provides the fit statistics for the top five qualifying distributions (where the distributions are ranked by BIC). The data shown in Figure 31 match those in Figure 30. The values appear different due to the different scaling. As noted in the submission, in order to focus on the best fitting distributions and maintain internal consistency with the selected base-case RFS distributions, only the top five best-fitting distributions which did not cross (i.e. OS < RFS throughout the model projection) the RFS unrestricted Gompertz distribution were considered.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

B18. Company submission, section B.3.3. page 103. Please clarify why OS curves which crossed RFS were excluded from further consideration when an alternative explanation of why the curves were crossing was that people were not relapsing in this period of the extrapolation? Please also clarify why, given this argument, the model still applies a logical consistency constraint which minimises RFS when the hazard exceeds that of the OS survivor function. Please also clarify the logic of this model selection criterion given that the base case PSA includes a proportion of samples where the RFS and OS curves cross.

While we recognise that it is possible for the RFS curve to meet the OS curve, we believe that this is unlikely, as it would suggest that no patients who failed to achieve response or who did achieve response and relapse would achieve a long-term cure. In the previous evaluation of blinatumomab for the treatment of R/R ALL (NICE TA450), it was confirmed by clinical experts that patients with R/R ALL could potentially be cured. Additionally, incorporation of instances in which the RFS curve meets the OS curve results in sudden changes in the hazard for RFS. While we believe that the inclusion of this criteria in the selection of curves for OS is not unreasonable, and use this in selection of the base case distributions, we include the logical consistency constraint in the model to permit consideration of combinations of RFS and OS that do not meet this criterion.

With respect to the PSA, while it is true that in some instances the curves do cross, the selection of a set of curves that do not cross in the base case reduces the likelihood of this occurrence in the PSA.

HRQoL

B19. PRIORITY. Company submission, section B.3.4 (page 116). Please provide further details regarding the GLM/GEE model:

a. Were other statistical model forms considered?

GLM/GEE regression was the only analytical approach considered, as this approach has been used and accepted in prior submissions to NICE.

b. Was there any control for clustering?

The GEE modelling approach controls for clustering (i.e., correlation of utility assessments within patients).

c. What was the distribution family for the data and link function of the GLM/GEE?

An identity link, normal error distribution, and exchangeable correlation structure was employed.

d. Please clarify why HSCT status was not included as a covariate in the GLM/GEE?

HSCT status was not included as a covariate as there were no utility assessments post HSCT in BLAST.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

B20. PRIORITY. Company submission, section B.3.4 (page 116). Please provide further details regarding the post-relapse EQ-5D estimates from BLAST. This should be presented as a table which includes the number of observations and the mean utility post-relapse at each timepoint.

In BLAST, there were a total of 8 post-relapse utility assessments. Of these, 6 assessments were conducted on the day of relapse, 1 on 22 days after relapse, and 1 on 30 days after relapse. The mean (SD) utility value for the 8 post-relapse assessments was 0.819 (0.276). Given the small number of post-relapse assessments, we do not believe it is appropriate to report mean values by timepoint.

B21. Company submission, section B.3.4, Table 47 (page 120). Please clarify why the age-adjusted utility formula published by Ara and Brazier (Value in Health, 13(5), Figure 2) was not used.

The alternative source for UK general population utility values referred to by the ERG likely represents a more robust source of data for these utilities as it is based on a larger and more recent sample. Utility values from the Ara and Brazier study are generally slightly higher than those based on the Kind report used in the original submission and consequently yields a slightly more favourable ICER (£27,938 (Ara) vs. £28,524 [Kind]). Applying the same utility data to the alternative base case (ie. blin as salvage) the ICER is £16,876 per QALY.

This calculation has been updated in the revised model included alongside our response.

Costs

B22. PRIORITY. Company submission, section B.3.5.4, page 128. Please clarify with details why the online survey responses demonstrated that some of the evidence survey respondents did not understand the exercise. Please clarify why only 2 experts were used to estimate health care resource use.

The HRU data for patients in first haematological CR with/without MRD was collected during a clinician survey study conducted in the EU5 countries in 2016 – 2017. The study was composed of two phases:

1. Pilot study: Two physicians from each of the five countries were recruited in this phase. Each selected physician completed a web-based questionnaire and a short telephone interview to collect feedback about the presentation and ease of use of the questionnaire which provided the clinicians more opportunities to better understand the questionnaire and minimise potential misinterpretation. The pilot phase interviews were conducted in English and, based on this feedback, the questionnaire was modified to improve ease of use and clarity. The pilot study HRU section questions were based on a 6-month average time period
2. Main study: A total of 103 physicians were recruited, 20 were from the UK. Each eligible physician completed a standardised web-based questionnaire about MRD testing in the treatment of adult patients with B-precursor ALL. The main study HRU section questions were based on a longer 12-month average time period which could potentially decrease reliability.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

In the main study, we found extremely unrealistic values of HRU reported, which led us to believe that there could have been a lack of understanding of the questionnaire. For example, as reported in Table 11 below (UK subset of pilot study), reported values from almost half of the surveyed clinicians were illogical, leading to doubt as to the reliability of the results.

Although this represents only one question, it highlights our concern that the format of the study may have led to clinicians not appropriately understanding the questions and thus reporting results that are not reflective of UK clinical practice. As a result – and as reported in the submission – we used estimates of health care resource from the in-depth interviews conducted in the pilot study as this was considered to provide a more robust estimate.

Table 11. Number of hospital admissions and duration of one hospital stay (for the last 12 months) for patients in CR1 (extreme values highlighted)

Respondent ID	Average # of hospitalisations	Average # of days per hospitalisation	Average total # of days in hospital in one year (calculated)
1	2	5	10
2	6	10	60
3	**75**	**25**	**1875**
4	**20**	**28**	**560**
5	1	2	2
6	**1**	**0**	**0**
7	3	12	36
8	2	10	20
9	1	7	7
10	**85**	**5**	**425**
11	**100**	**28**	**2800**
12	1	5	5
13	8	14	112
14	1	10	10
15	**60**	**70**	**4200**
16	**3**	**135**	**405**
17	**85**	**7**	**595**
18	**6**	**60**	**360**
19	6	10	60
20	**30**	**30**	**900**

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

B23. Company submission, section B.3.5, Table 49, page 122-123. Please provide the source for the cost of a pump applied in the model.

The cost of the pump used in the model was calculated based on input from UK oncology nurses considering the pump to be a BodyGuard 323™ Ambulatory Infusion Pump. Specific inputs for the pump cost, maintenance costs and consumables were sourced directly from the supplier and are consistent with the approach taken for the NICE TA450 appraisal.

B24. Company submission, section B.3.5, page 122. Please clarify whether the economic analysis accounts for the days that a pump was not allocated to a person receiving blinatumomab within its 5-year lifespan.

In the model, the costs of the pump are calculated based on the prorated daily cost and the number of days the pump was used. This approach is premised on the assumption that the pump could be used by another patient on the days that the pump is not being used by the patient receiving blinatumomab. We recognise that this approach might underestimate the cost of the pump if the patients hold on to the pump in between cycles; however, any such underestimation would have a negligible impact on the ICER given the relatively low cost of the pump.

This calculation has been updated in the revised model included alongside our response.

B25. Company submission, section B.3.2.2, page 89. The model assumes there are no disease-related costs after 5 years. Please justify this assumption

This assumption is aligned with UK clinical expert opinion that consistently indicated that ALL-related costs (excluding follow-up costs associated with HSCT conducted previously) would be zero at five years – this therefore reflects the timepoint at which clinicians consider patients who remain alive would be cured.

B26. Company submission, section B.3.5.4, page 126. Please provide the filename in the reference pack, a PDF or clear web link to reference 105 of the CS.

The aforementioned reference has been included alongside this response.

B27. Company submission, section B.3.5.4, page 127. Given the differences between the populations under appraisal, please comment on the appropriateness of taking the costs of salvage therapy from the previous STA of blinotumomab for people with previously treated B-precursor acute lymphoblastic leukaemia.

In the base-case analysis, it was assumed that all patients would receive SoC salvage therapy upon relapse – this assumption is aligned with the expected management of these patients in clinical practice. The cost of salvage therapy was based on medication and administration costs for FLAG-IDA and estimated using an economic model submitted as a part of TA450 (blinatumomab for previously treated B-precursor acute lymphoblastic leukaemia). FLAG-IDA was considered to be the most appropriate regimen to capture in this evaluation of MRD+ as it best reflects standard of care for patients with relapsed B-cell precursor ALL.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

To retain internal consistency and ensure alignment with the modelled population (i.e. Ph-, CR1), calculations were estimated using the 'no prior salvage' subgroup. These populations are considered to be similar in terms of patient characteristics and how they are managed in clinical practice.

B28. Company submission, section B.3.5.4, page 127. Please provide a source for the 37% of people who receive first-line salvage therapy who go on to receive second-line salvage therapy. What was the time to second-line salvage therapy (for those who received this treatment) used in this model.

In the base case, it was assumed that all patients would receive SoC salvage therapy upon relapse. The cost of SoC salvage therapy was estimated using an economic model used in the manufacturer's submission in response to STA of blinatumomab for previously treated B-precursor ALL based on the TOWER trial (STA 1804). This model was used to calculate the costs of first and second salvage therapy for the no prior salvage therapy subgroup of patients randomised to SoC assuming that all patients who relapse would receive first-line salvage therapy, that 37.0% of patients who relapse after first-line salvage therapy would receive second-line salvage therapy, and that the cost per course of salvage therapy is £16,175, based on medication and administration costs for FLAG-IDA. The proportion of patients experiencing relapse who received salvage therapy was from the TOWER trial. The distribution of time to salvage therapy was not calculated in this model. However, the mean time in the initial (pre-response assessment) and response state was 8 months.

B29. Company submission, section B.3.5.4, page 127. Please comment on the appropriateness of assuming that the same chemotherapy regimen is given upon a second relapse in the population who receive salvage therapy?

The administration and medication costs of subsequent salvage therapies are assumed to be the same as initial salvage therapy costs and is consistent with the approach taken for the appraisal of blinatumomab for R/R ALL (NICE TA450). Given the similarities of salvage regimens (not including innovative/experimental therapies) and the small number of patients receiving subsequent treatment, the impact of this assumption is minimal.

Model

B30. PRIORITY. Company submission, section B.3.2, page 86. Given the need to track HSCT in order to estimate both costs and health outcomes, please clarify why the model has been implemented as a partitioned survival model rather than a state transition model.

The reasons for using the partitioned survival model are described on page 87 of the submission. While we did incorporate the impact on HSCT on costs and quality of life, we did not explicitly model the impact of HSCT on survival. Given the small numbers of patients who did not undergo a transplant in BLAST, and limited access* to data from the historical control study, it was not feasible to develop a model with states defined on HSCT.

**As a point of clarity, Amgen does not have ownership of the data collected by the principle investigators (PIs) in the historical control study as this remains with the PIs*

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

B31. PRIORITY. Model. Please explain how to use the model to estimate the total number of HSCTs pre-relapse and post-relapse over the lifetime of the model cohort.

The numbers of pre- and post-relapse HSCTs can be obtained from cells GU133 and HL133, respectively, of the Blin Calc and SoC Calc sheets.

B32. PRIORITY. Model, "Blin Calc" worksheet, GQ9:HW129 & Model, "SOC Calc" worksheet, GQ9:HW129 Please clarify the logic of all calculations used to approximate HSCT receipt and it's associated treatment over time.

These cells are used to project the incidence and costs of pre- and post-relapse HSCT in the model and are effectively a nested Markov cohort model with a six month cycle duration and states defined in the occurrence of HSCT and 6 tunnel states for time since HSCT. As noted in the submission, it was assumed that patients with pre-relapse HSCT would not relapse until all patients without pre-relapse HSCT have relapsed. Under this assumption, all relapses occurring prior to the point at which the RFS and cumulative HSCT curves cross are assumed to be among patients with no prior HSCT while all those occurring after that point are assumed to be amongst those with prior relapse. Also, for discounting purposes, the cost of post-relapse HSCT was assumed to occur at the time of relapse. Thus, in the model, the probability of post-relapse HSCT is calculated by combining the proportion of patients relapsing each 6 months (the difference in successive values in column HJ) and the probability of HSCT given relapse, conditioned on whether RFS (column HJ) is greater than the cumulative incidence of pre-relapse HSCT (column HG).

B33. PRIORITY. Model, "Cost Inputs" worksheet, F109:F110. Please clarify the time period for which the probability of receiving HSCT upon relapse (stratified by prior HSCT status) was calculated.

As noted in the submission and above in response to Question B32, lacking information on the timing of post-relapse HSCT, for the purpose of discounting, post-relapse HSCT was assumed to occur at the time of relapse. This may result in an overestimate of the discounted cost of post-relapse transplant, which may impart a slight bias in favour of blinatumomab (since the incidence of post-relapse HSCT is higher with SoC). It should be noted that the occurrence of post-relapse HSCT is not limited to 60 months. To be consistent with the approach for including other ALL related costs, it may be more appropriate to limit the occurrence of post-relapse HSCT to 60 months. Employing this restriction yields a slightly more favourable ICER of £28,327 per QALY. Applying the restriction to the alternative base case (ie. blin as salvage) results in an ICER of £17,120 per QALY.

This calculation has been updated in the revised model included alongside our response.

B34. Company submission, section B.3.6, Table 55, page 131. Please clarify why the PSA does not include any uncertainty around MRD response for the SoC group. Please also clarify why the PSA does not include any uncertainty around the proportion of RFS events that are deaths.

We did not sample the probability of MRD response for patients receiving SoC as this parameter was based on assumption and we had no estimates of its distributional properties. However, we

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

provided scenario analyses assuming different probabilities of MRD response for patients receiving SOC and the ICER was not sensitive to this parameter.

The model does sample the proportion of RFS events that are deaths (see cells b10-J12 in the PSA Input Sheet).

B35. Company submission, section B.2.6.1, page 59 and company submission, section B.2.9.2, Table 27, pages 67-68. Given the importance attributed to the 100-day mortality associated with HSCT in the clinical section of the CS, please clarify why this effect was not explicitly included in the health economic model.

Although 100-day mortality after HSCT was lower in lower in BLAST (7%) than published estimates (>25%)[11], this was not modelled explicitly, as treatment effects on mortality post-transplant were captured implicitly in treatment effects on OS.

B36. Company submission, section B.3.5.4, page 126. Please clarify whether the 38.4% of people in the SoC arm receiving HSCT after 4 years refers only to the post-matching population.

Due to restrictions on access to data from Study 20120148, we were unable to conduct a propensity-matched analysis of the percent of patients undergoing HSCT. The 38.4% used in the model for SoC is based on the unmatched population of Study 20120148.

B37. Company submission, section 3.5.4, page 126. Why were data on post-relapse HSCT not available in either BLAST or the historical control study?

Data capture on post-relapse HSCT was unfortunately not included in the respective protocols thus we are unable to provide further data here.

B38. Company submission, section 3.5.4, page 126. If you haven't done so in response to a previous question, please clarify how the probabilities that a patient received HSCT upon relapse were calculated from the data in BLAST and Clinicaltrials.gov: NCT02003612.

As noted in the submission, data on the probability of HSCT after relapse were unavailable from BLAST or Study 20120148. Accordingly, we used data from Protocol 20120310 (NCT02003612) to obtain estimates of the probability of undergoing HSCT post-relapse conditioned on whether the patient had undergoing HSCT pre-relapse. Data on the frequency of HSCT for patients who had not received prior salvage in Protocol 20120310 was available by age (<35 vs. ≥35 years) and for patients with and without prior HSCT. We therefore weighted the age-specific data using the age distribution of relapsing patients in BLAST to obtain estimates of the probability of HSCT in patients with and without prior HSCT.

B39. Company submission, page 127. Please clarify how the £16,175 per line of salvage chemotherapy was calculated. In the publicly available CS associated with the referenced FAD (Amgen Ltd, Blinatumomab for previously treated B precursor acute lymphoblastic leukaemia: Company evidence submission, page 164, Table 5-15) only a cost of £13,438 per cycle of FLAG-IDA was available.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

The £16,175 cost per line of salvage implemented in the model is consistent with the no prior salvage subgroup evaluated in TA450. This population was used as they were most representative of patients from BLAST and was consistent with the treatment pathway for patients in CR1.

B40. Model, "Blin Calc" worksheet, cells EQ9:EQ12. Please clarify why the daily pro-rated pump cost and the annual maintenance cost of the pump divided by 365 are applied to the number of outpatient treatment days (cells EC9:EC12) rather than the number of days within a treatment cycle.

We calculate the costs of the pump assuming that the pump could be used by another patient on the days on which the patient receiving blinatumomab was not using it. We recognise, however, that this may yield an underestimate of the number of days (see response to **B24**). Pump costs are not applied to inpatient days as the model assumes that the costs associated with blinatumomab administration is captured by the inpatient costs applied. This assumption would have a negligible impact on the ICER given the relatively low cost of the pump.

B41. Model, "PSA Bootstrap Inputs" worksheet & Model, "PSA Results" worksheet, cell E4. Please clarify why fewer bootstrap samples of the parametric distribution parameters are included compared with the PSA samples (1,000 bootstrap samples versus 10,000 PSA samples).

The joint bootstrap distribution was generated with 1000 bootstrap estimates as it was felt that this number of samples would be sufficient to reasonably characterised the distributions of the parameters included in the joint distribution and the inclusion of additional bootstrap estimates would further increase the size and potentially slow the calculations of the model. 10,000 simulations were used for the PSA in order to yield relatively precise estimates of the percentiles of the distributions of the PSA outcomes. While there are only 1000 possible realisations from the bootstrap samples, there are numerous other parameters sampled in the PSA, so the total number of potential realisations far exceeds 10,000. We therefore do not believe that use of a bootstrap distribution with 1000 bootstrap estimates materially biases the estimates derived from the PSA.

B42. Model, "Blino calcs" worksheet, columns AM: AN. The (uplifted) general probability of death is greater than 1.0 for later ages. Please comment.

The ERG is correct that the adjusted general population mortality is greater than 100% at age 94+ for men and 97+ for women. As a consequence, survival beyond this age is set to zero. This is a consequence of applying the risk ratio as a scalar to the probability. While this may result in a slight overestimation of mortality and underestimation of life expectancy, the effect of this bias is not material.

B43. Model, "PSA Inputs" worksheet, I104:I122. Please clarify how the standard errors for the calibrated 6 monthly probabilities of receiving HSCT pre-relapse were calculated.

The SEs for the calibrated 6 month probabilities of HSCT are in error. The correct values 4.08% and 1.95%, which were calculated using the formula SE=sqrt(p*(1-p)/N), where N=73 and N=287 for SoC, respectively. The use of the corrected values does not materially impact the results of the PSA. We recognise that this approach to calculating the SEs of the calibrated values (based

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

on the overall N) may not be precise, but its use is not likely to materially impact the results of the PSA.

B44. Model, “Blino calcs” worksheet, columns AW and BA. The mortality probability does not change at the integer age. Please clarify.

In the model, the probability of death based on the general population morality is based on age rounded down to an integer. That is, for the first year of the model, the general population mortality probability corresponds to that for a 45-year-old person. The mortality probability in the second year pf the model, corresponds to that for a 46-year-old persons, and so on. Column AT, labelled "Age", is not used in the model.

EoL

B45. Company submission, section B.2.9.4, page 75 and Company submission, section B.2.13.3, Table 32, page 85. Please explain why the OS figures are quoted based on the matched population of the historical control rather than the unweighted data from this source?

The mean estimates of OS are based on the modelled survival projections and thus reflect the matched population of the historical control used in the base case analysis.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

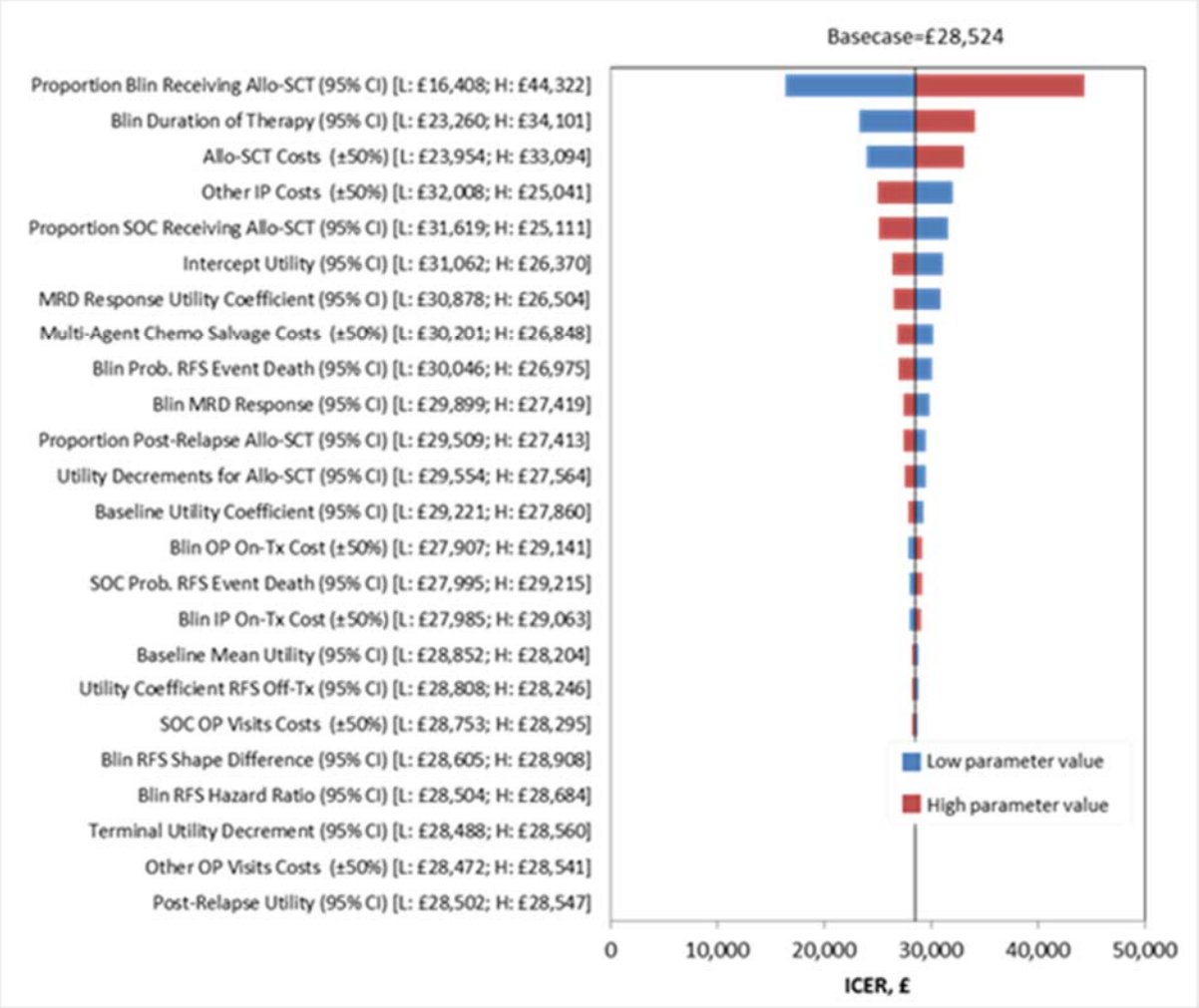
+44 (0)845 003 7780

Further Clarification Questions

1. **In Document B figure 45 on page 153 is not visible. This was the case in both the original and updated submissions. Please could you check this and provide figure 45?**

A tornado chart for the ICER for blinatumomab vs. SoC is shown in Figure 2 (Figure 45 in the original submission).

Figure 2. Tornado diagram of ICER of blinatumomab versus SoC



NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

2. **You have provided two versions of the clarification responses. One of these is a marked version. The second one is labelled clean which appears to be an unmarked version (i.e. contains confidential data which is not marked up). The two versions we require are a marked version (that we have) and a redacted document with all the confidential information blacked out. Please could you provided a redacted version of the clarification response document?**

Marked-up and redacted version of our responses have been now been provided.

3. **A8 additional query from ERG: It is acknowledged that there is an inconsistency between the results presented in the clinical effectiveness section and the cost effectiveness section. The clarification response states "We would be happy to provide an updated version of the calculations to confirm consistency of the results but have been unable to complete this in the timeframe – however, this can be provided if required". The ERG do not require this additional analysis, however could the company please provide a table outlining the covariate balance before and after propensity score adjustment using ordinary ATT weights (as in table 29, but using ordinary rather than the stabilised weights). This is so we have the clinical-effectiveness evidence that is consistent with the data that has been used in model.**

Further to the response provided in QA8, the following table presents the covariate balance before and after propensity score adjustment using ordinary (ie. non-stabilised) ATT weights.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

Table 12: Difference in Patient Characteristics by Treatment Group - After IPTW Adjustment

Characteristic	Unweighted				IPTW			
Mean (SD)/n (%)	Control [redacted]	Blinatumomab [redacted]	Standard Difference	P-value	Control [redacted]	Blinatumomab [redacted]	Standard Difference	P-value
Age at primary diagnosis (years)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gender (Female)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Country (Not Germany)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
MRD at Baseline (recoded)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Time from diagnosis to baseline (months)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
WBC at diagnosis (>30,000/mm^3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
WBC at diagnosis (continuous, log10)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
T411mll4 mutation (Yes)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Prior chemotherapy (GMALL)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

References

1. Amgen I. Data on File: Clinical Overview (Blinatumomab). TBC.
2. European Network for Health Technology Assessment. Endpoints used in relative effectiveness assessments of pharmaceuticals. 2013.
3. European Medicines Agency. European Public Assessment Report: Summary of the risk management plan (RMP) for Blincyto (blinatumomab). 2015.
4. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika 1983;70:41-55.
5. Yue LQ. Statistical and regulatory issues with the application of propensity score analysis to nonrandomized medical device clinical studies. J Biopharm Stat 2007;17:1-13; discussion 15-7, 19-21, 23-7 passim.
6. Levenson MS, Yue LQ. Regulatory issues of propensity score methodology application to drug and device safety studies. J Biopharm Stat 2013;23:110-21.
7. Austin PC. The performance of different propensity score methods for estimating marginal hazard ratios. Stat Med 2013;32:2837-49.
8. Austin PC, Grootendorst P, Normand SL, et al. Conditioning on the propensity score can result in biased estimation of common measures of treatment effect: a Monte Carlo study. Stat Med 2007;26:754-68.
9. Barnow B, Cain G, Goldberger A. Issues in the Analysis of Selectivity Bias. Evaluation Studies 1980;5:42-59.
10. Cole SR, Hernan MA. Adjusted survival curves with inverse probability weights. Comput Methods Programs Biomed 2004;75:45-9.
11. Bishop, M. R., et al. "Long-term outcomes of adults with acute lymphoblastic leukemia after autologous or unrelated donor bone marrow transplantation: a comparative analysis by the National Marrow Donor Program and Center for International Blood and Marrow Transplant Research." Bone marrow transplantation 41.7 (2008): 635-642.

Single technology appraisal

Blinatumomab for acute lymphoblastic leukaemia for people with minimal residual disease activity in remission [ID1036]

Clarification on response to ERG Questions

June 2018

Ahead of the NICE AC Meeting for blinatumomab [ID1036] we wanted to take the opportunity to proactively address an error in our response to the ERG clarification questions. This error, which has recently come to our attention, relates to an additional query raised on Question A8 and concerns the implementation of the ATT weights in the economic evaluation. This clarification has no material impact on the results of the economic evaluation.

We have provided a corrected response below, using our original response with amendments highlighted, to supersede our previous response.

A8 Additional Query. CS Table 29 page 73. The baseline characteristics of the SOC and Blinatumomab groups are presented before and after propensity score adjustment with ATT weights. Could you please confirm that the figures presented in the table, and specifically the sample size (N=174.3 control, N=78.5 blinatumomab) are correct? The formulae presented in Appendix L page 219 indicate that the under the ATT assumption the treatment arm should not be weighted (weight=1). The curves presented in the cost effectiveness section, Figure 22, page 97, indicate N=155.1 control, N=73 blinatumomab. This relates to question **A8** in the original clarification letter.

There are two issues that we would like to address/correct in this response: **1)** Reporting of sample size for stabalised IPTW analysis; **2)** Application of weights in cost-effectiveness analysis.

Reporting of sample size in stabilised IPTW analysis

The sample size presented in Table 29 of the submission were unfortunately reported incorrectly – please find updated Table below with the correct sample size reported. It is important to note that these values reflect the baseline characteristics after the propensity score adjustment using the *stablised* IPTW (sIPTW) and as a result the weighting for the blinatumomab arm is not equal to one.

Stabilised IPTW can be conducted to reduce the potential instability caused by very large weights' and was applied for this propensity score analysis. In order to calculate stabilised weights, the IPTW is multiplied by the marginal probability of receiving the actual treatment received (Cole and Hernan, 2004).

$$sw_{j,k} = w_i \frac{n_k}{\sum_{k=1}^{2} n_k}$$

Where $sw_{j,k}$ represents the sIPTW for the jth subject from treatment k and n_k represents the sample size for treatment k. This results in non-integer values for both blinatumomab and SoC in this analysis.

Characteristic	Unweighted				Stabilised IPTW			
Mean (SD)/n (%)	Control [redacted]	Blinatumomab [redacted]	Standard Difference	P-value	Control [redacted]	Blinatumomab [redacted]	Standard Difference	P-value
Age at primary diagnosis (years)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gender (Female)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Country (Not Germany)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
MRD at Baseline (recoded)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Time from diagnosis to baseline (months)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
WBC at diagnosis (>30,000/mm^3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
WBC at diagnosis (continuous, log10)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
T411mll4 mutation (Yes)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Prior chemotherapy (GMALL)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Application of weights in Cost-effectiveness analysis

In the cost-effectiveness analysis, the ATT propensity score analysis was applied using an alternative method of ~~non~~-stabilised ATT weights. Unfortunately, this was done in error and lacks consistency with the presentation of clinical results.

The calculations of the stabilised weights are summarised in the Table below:

Stabilized ATT (Method 1 – clinical effectiveness)		**Stabilized ATT (Method 2 – cost-effectiveness)**	
BLIN weight	SOC weight	BLIN weight	SOC weight
p	$\frac{pi}{1-pi} \times (1-p)$	1	$\frac{pi}{1-pi} \times \frac{1-p}{p}$

However, we are confident that the application of this alternative method of stabilized ATT weights would have no impact on the analyses conducted for cost-effectiveness since the relevant size of the two groups are the same (73:155.1=0.47; 20.9:44.4=0.47). We would be happy to provide an updated version of the calculations to confirm consistency of the results but have been unable to complete this in the timeframe – however, this can be provided if required.

Furthermore, the ERG concluded that the standard weights (ie. unstabilised) in the economic evaluation were appropriate. Although the unstabilised ATT weights results in a difference in the relative sample size between blinatumomab and SOC, the shapes of the OS/RFS curves remain identical. As such, when implementing these in the economic evaluation, the ICER does not meaningfully change from the base case ICER presented in our submission (delta ~£100/QALY). We would be happy to provide more information on this point if required.

NICE National Institute for Health and Care Excellence

Patient organisation submission

Blinatumomab for treating minimal residual B-precursor acute lymphoblastic leukaemia in remission [ID1036]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	[redacted]

2. Name of organisation	Leukaemia Care
3. Job title or position	[redacted]
4a. Brief description of the organisation (including who funds it). How many members does it have?	Leukaemia Care is a national blood cancer charity, first registered with the Charity Commission in 1969. We work to ensure that everybody has the right information, advice and support. Our key services include: Freephone helpline, Nurse Advisor, LiveChat, Nationwide Support Groups, Conferences, Campaigning and Advocacy, Buddy Support and Patient Booklets. Over 85% of our funding comes from our own fundraising activities and those of our volunteers. This includes a wide range of activities – such as legacies, community events, marathons, recycling campaigns etc. Leukaemia CARE also receives funding from a wide range of pharmaceutical companies, but in total those funds do not exceed 15% of our total annual income. Any funds received from the pharmaceutical industry are in accordance with the ABPI Code of Practice and the Leukaemia Care Code of Practice, our voluntary commitment that governs how we work with, and accept funding from, the pharmaceutical industry: www.leukaemiacare.org.uk/resources/code-of-practice
4b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	N/A

5. How did you gather information about the experiences of patients and carers to include in your submission?	Information primarily gathered through Leukaemia Care patient experience survey – 'Living with Leukaemia' (www.leukaemiacare.org.uk/living-with-leukaemia). The survey was run from September to December 2016, as a follow up to NHS England's annual Cancer Patient Experience Survey (NCPES). The Leukaemia Care survey involved 85 questions and had responses from 2519 blood cancer patients, including 151 acute lymphoblastic leukaemia (ALL) patients. The results of this survey have been used to inform our submission. Additionally, we have gathered information through our helpline, support groups, communication with our membership and one to one discussion with patients. We also work closely with other patient groups and share expertise.
Living with the condition	
6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	Acute lymphoblastic leukaemia (ALL) is a rare and rapidly progressing form of leukaemia. In 2014, there were 758 new cases of acute lymphoblastic leukaemia in the UK. Approximately 60% of these cases were diagnosed in children and teenagers. Most of the remaining 300 cases were diagnosed in adults over the age of 50. Symptoms experienced prior to diagnosis include fatigue (69%); feeling weak or breathless (61%), fever or night sweats (36%), bruising or bleeding (31%), pain in bones or joints (28%), unexplained weight loss (26%), sleeping problems (26%) and swollen lymph nodes (22%). Due to the rapidly progressing nature of the condition, 63% of patients had experienced symptoms for less than a month before visiting their GP.

	The NCIN/NCRAS routes to diagnosis report shows that 64% of ALL patients are diagnosed via emergency presentation (of which 42% were A&E, 27% emergency GP referral, 5% inpatient emergency and 26% outpatient emergency). This compares to a cancer average of 22% and is the highest of any cancer type in the report. The rapidly progressing nature of the condition means that 86% of ALL patients start treatment within a week of diagnosis. Being diagnosed with ALL can also have a huge emotional impact, prompting patients (and their families) to experience feelings of disbelief, denial, anger, fear, blame, guilt, isolation and depression. In our survey, 60% of ALL patients reported that they have felt depressed or anxious more often since their diagnosis. The emotional impact does not affect the patient in isolation. A diagnosis can place huge emotional strain on families and friends, many of whom may be affected. As such, improvements in a patients' treatment and quality of life will also have a wider impact on the lives of their family and friends. The most common symptoms encountered by patients since their diagnosis are fatigue (76%), feeling weak or breathless (54%), sleeping problems (53%), nausea or vomiting (45%), memory loss or loss of concentration (44%), tingling or numbness in extremities (44%), bone or joint pain (38%), bleeding or bruising (38%) and infections (36%). ALL also has a much wider practical impact, with 64% of ALL patients experiencing pain as a direct result of their condition (30% occasionally, 24% regularly and 9% constantly). Additionally, 62% of ALL patients have difficulty moving around (sometimes 32%, often 18% and always 11%) and 65% of ALL patients have difficulty performing some of their daily routines, such as cooking or cleaning. Another 48% reported

	that they have problems taking care of themselves (sometimes 31%, often 11% or unable to self-care at all 6%). Of those in work or education before their diagnosis, 70% have been impacted (31% reduced hours, 39% no longer able to work or continue education). Consequently, 60% of ALL patients reported a negative financial impact as a result of having cancer (increased costs or reduced income). Five-year survival outcomes vary greatly by age, from over 90% in the under 14s, almost 70% in those aged 15-24, less than 40% in those aged 25-64 and less than 15% in those aged 64 or older. As such, the prognosis for adult patients with ALL is extremely poor.
Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	Whilst highly toxic chemotherapies have high response rates (80-90%), nearly half of patients will eventually relapse. Having identified MRD activity following induction (around a third of patients), these patients would be categorised as high-risk and likely proceed to transplant (if eligible), but with poor outcomes. Most patients (80%) will relapse and die within two years, likely in around 6-9 months. Around 20% will be potentially 'cured' following transplant and have long-term life expectancy. In our survey, 41% of ALL patients said they don't think there are currently enough treatment options available for ALL on the NHS. The most common side effects of treatment reported by ALL patients were fatigue (74%), neutropenia (44%), nausea or vomiting (44%), hair loss (42%), muscle or joint pain (41%), sore mouth (40%), sleeping

	problems (38%), loss of concentration or memory (37%), diarrhoea (32%), bone and joint pain (32%). The combined impact of these side effects was rated by 52% as having had a large impact, with 51% of patients hospitalised as a result of side effects.
8. Is there an unmet need for patients with this condition?	Yes – in this setting there is an urgent need for access to treatments that can induce MRD negativity, prevent relapse and improve survival outcomes.
Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	Blinatumomab is extremely effective (78%) in inducing MRD negativity in patients, usually within the first treatment cycle (98%). In patients with MRD activity following intensive therapy, inducing MRD negativity with blinatumomab has the potential to improve outcomes, usually through acting as a bridging therapy to a subsequent allo SCT (with improved outcomes). In our patient survey, 91% of patients said they would welcome a treatment that would enable them to have an SCT. Without access to blinatumomab in this setting, nearly half of patients will eventually relapse – many following an SCT. Patients report relapse as one of the hardest parts of their journey – with patients who reported experiencing a relapse more likely to say they felt constantly depressed or anxious (18%

	relapsed ALL v 4% ALL without relapse). Blinatumomab has the potential to be administered as an outpatient (the most popular setting in our survey) and through an IV infusion (the second most popular option in our survey for ALL patients, after an oral tablet).
Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	Common side effects include fever, headaches, tremors, chills, fatigue, nausea and vomiting. These are not unusual for ALL treatment and indeed blinatumomab is generally deemed manageable/tolerable. In our recent survey, 76% of ALL patients reported that they would be willing to experience additional side effects for a more effective treatment.
Patient population	
11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	Yes – those who would subsequently relapse without access to blinatumomab. The strongest predictive marker for this is MRD activity.

Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	
Other issues	
13. Are there any other issues that you would like the committee to consider?	

Key messages

15. In up to 5 bullet points, please summarise the key messages of your submission:

- Acute lymphoblastic leukaemia (ALL) is a rare and aggressive disease, with a significant symptom burden and mortality. It progresses rapidly, resulting in high numbers of emergency referrals and a need to rapidly start treatment. This can place a huge emotional strain on patients and their families.
- Patients assessed to have MRD activity following induction treatment would be considered high-risk, with poor survival (a matter of months). There is an urgent need for access to treatments that can induce MRD negativity, prevent relapse and improve survival outcomes.
- Blinatumomab is extremely effective (78%) in inducing MRD negativity in patients, usually within the first treatment cycle (98%). In patients with MRD activity following intensive therapy, blinatumomab has the potential to improve outcomes, usually through acting as a bridging therapy to a subsequent allo SCT. In our patient survey, 91% of patients said they would welcome a treatment that would enable them to have an SCT.
- Without access to blinatumomab in this setting, nearly half of patients will eventually relapse – many following an SCT. Patients report relapse as one of the hardest parts of their journey – with patients who reported experiencing a relapse more likely to say they felt constantly depressed or anxious (18% relapsed ALL v 4% ALL without relapse).

Thank you for your time. Please log in to your NICE Docs account to upload your completed submission.

NICE National Institute for Health and Care Excellence

Professional organisation submission

Blinatumomab for acute lymphoblastic leukaemia for people with minimal residual disease activity in remission [ID1036]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	[redacted]
2. Name of organisation	**Royal College of Pathologists**

3. Job title or position	[redacted] [redacted]
4. Are you (please tick all that apply):	☒ an employee or representative of a healthcare professional organisation that represents clinicians? ☒ a specialist in the treatment of people with this condition? ☐ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5a. Brief description of the organisation (including who funds it).	**The Royal College of Pathologists is a professional membership organisation committed to promoting excellence in the practice of pathology.**
5b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	**No**
The aim of treatment for this condition	
6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition,	The overall aim in the treatment of Acute Lymphoblastic Leukaemia is cure. Broadly speaking this is achieved in two main phases. Firstly, remission induction, where the aim is to clear the majority of the leukaemia. Secondly consolidation, where the aim is to consolidate the remission so that any small amounts of remaining leukaemia are treated to reduce the risk of relapse. This is usually attempted with either chemotherapy treatment or donor stem cell transplantation.

or prevent progression or disability.)	
7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	A clinically significant response to treatment can be assessed either by depth of remission or duration of remission. Conventional assessment has been by counting the number of leukaemia cells observed under a microscope on a bone marrow smear. A patient is deemed to be in remission when the percentage of leukaemia cells is less than 5%. However this method lacks sensitivity. More recently bone marrow assessment by Minimal Residual Disease (MRD) has been found to be more accurate and enable detection of much lower levels of leukaemia. It also correlates strongly with outcome, i.e. those patients who go into a deeper remission either are less likely to relapse or if they do they relapse later.
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Definitely yes. Those patients who are MRD positive after chemotherapy have a poor outlook. Further chemotherapy usually does not result in a durable response and donor stem cell transplantation is often ineffective. Blinatumomab treatment has excellent results in this situation of residual MRD positivity.
What is the expected place of the technology in current practice?	
9. How is the condition currently treated in the NHS?	Currently the options in a patient who remains MRD positive following on from induction chemotherapy are further chemotherapy or donor stem cell transplantation. Neither of these options are satisfactory.
• Are any clinical guidelines used in the treatment of the	Currently there are no National Guidelines for treatment of this condition. The most widely used Guidelines are the American NCCN guidelines,

condition, and if so, which?	https://www.nccn.org/professionals/physician_gls/pdf/all.pdf These recommend Blinatumomab for the treatment of persistent or late clearance of MRD.
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	The current treatment pathway for ALL is well defined. Most centres treating patients 25 years and over would use the UKALL14 Trial Protocol either as part of the trial or off trial. Where patients are treated off trial the chemotherapy backbone used is fairly consistent but there are some differences in the use of donor stem cell transplantation and the details of how this is performed (e.g. reduced intensity conditioning or full intensity conditioning regimens).
• What impact would the technology have on the current pathway of care?	This technology assessment is looking at the use of Blinatumomab in the context of minimal residual disease eradication. This is an area of unmet need without a standard treatment. The results from a Phase 2 study show that this treatment is safe and effective with 80% of patients treated becoming MRD negative.
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Current use is variable. A limited number of centres have used it in the context of clinical trials and more have used it through a company compassionate use scheme.
• How does healthcare resource use differ between the technology and current care?	N/A

• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	This drug is ideally used in the setting of residual minimal disease which is the focus of this assessment. Whilst nearly all haematology units will be familiar with the infusion of monoclonal antibodies this technology does require both familiarity with managing the potential toxicities, specialist equipment (Specific infusion pumps) and preferably an ambulatory care programme. With this it can largely be given as an outpatient. In my opinion this treatment should only be given by Allogeneic transplant units.
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	Currently most allogeneic transplant unit should meet the above requirements. Other units might need significant equipment and training.
11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes I definitely think this will provide clinically meaningful benefits to patients e.g. less patients requiring second line chemotherapy treatment and more patients ultimately being cured of their leukaemia.
• Do you expect the technology to increase length of life more than current care?	Definitely yes. Acute Lymphoblastic Leukaemia is also frequently seen in younger adults who are more likely to return to 'tax paying status' if successfully treated,
• Do you expect the technology to increase health-related quality of life more than current care?	Definitely yes. The treatment is very well tolerated especially compared with second line chemotherapy treatment.

12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	No.
The use of the technology	
13. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	The technology is initially more difficult for healthcare workers unless they are experienced and have ambulatory care facilities. Now I have used this drug regularly over three years it is easier to administer than intensive chemotherapy. For patients it is much better tolerated than intensive chemotherapy and they have the benefit of less time in hospital compared with intensive chemotherapy.

14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Clear rules should include, 1. Indications for its use i.e. treatment of MRD positivity following induction chemotherapy and treatment of MRD positivity either as a bridge to donor stem cell transplantation or following donor stem cell transplantation. 2. Stopping rules. Repeat MRD assessment on bone marrow after 1 and/or 2 cycles of treatment. 3. Maximum cycles to be used in those responding =4.
15. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	Not that I am aware of.
16. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related	Yes.

benefits and how might it improve the way that current need is met?	
• Is the technology a 'step-change' in the management of the condition?	Yes. There is currently no good treatment of MRD positivity. Options are to accept it and wait for the patient to relapse or to treat it as frank relapse with second line chemotherapy treatment.
• Does the use of the technology address any particular unmet need of the patient population?	Yes. It is a chemotherapy free treatment in those where chemotherapy is not working.
17. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	The two main side effects are infusional related and neurotoxicity. The first is very manageable and the second is considered transient and fully reversible.
Sources of evidence	
18. Do the clinical trials on the technology reflect current UK clinical practice?	Yes.

• If not, how could the results be extrapolated to the UK setting?	N/A
• What, in your view, are the most important outcomes, and were they measured in the trials?	In the treatment of MRD positivity the most important outcome is to attain MRD negativity and this was achieved in 80% of patients
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	MRD is a surrogate outcome. Longer follow up of patients treated on the trial showed durable responses both in those who underwent subsequent donor stem cell transplantation and in those who did not.
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	No.
19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No.
20. Are you aware of any new evidence for the comparator	No comparators

treatment(s) since the publication of NICE technology appraisal guidance [TAXXX]?	
21. How do data on real-world experience compare with the trial data?	I do not have access to this but AMGEN have collected some limited data on their compassionate use programme.
Equality	
22a. Are there any potential equality issues that should be taken into account when considering this treatment?	No.
22b. Consider whether these issues are different from issues with current care and why.	
Topic-specific questions	
23.	FLAG +/-Idarubicin, Clofarabine based chemotherapy.

What combination chemotherapy regimens are typically used for the treatment of minimal residual disease? Would retreatment with blinatumomab be considered if relapsed occurred?	Often it is not treated. Good question. MRD recurrence is not the same as frank relapse. I would use Blinatumomab a second time to treat MRD if it was available e.g. before and after stem cell tranaplant.

Key messages

24. In up to 5 bullet points, please summarise the key messages of your submission.

- It is effective -80% conversion of MRD positive to negative
- It is safe
- Those successfully treated are often young and may go on to live long lives
- No real alternatives
- Reduces frank relapses and need for unpleasant and dangerous second line chemotherapy treatments

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

NICE National Institute for Health and Care Excellence

Clinical expert statement

Blinatumomab for treating minimal residual B-precursor acute lymphoblastic leukaemia in remission [ID1036]

Thank you for agreeing to give us your views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	**Dr N.J.Morley**
2. Name of organisation	**Sheffield Teaching Hospital NHS Foundation Trust**

3. Job title or position	**Consultant Haematology**
4. Are you (please tick all that apply):	☒ an employee or representative of a healthcare professional organisation that represents clinicians? ☒ a specialist in the treatment of people with this condition? ☐ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☒ yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ other (they didn't submit one, I don't know if they submitted one etc.)
6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission.)	☒ yes

The aim of treatment for this condition	
7. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	To improve cure rate, overall survival and duration of remission by eliminating minimal residual disease in patients with Acute Lymphoblastic Leukaemia (i.e treating very low levels of leukaemia cells that remain despite chemotherapy treatment).
8. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	Traditionally leukaemia has been assessed by counting leukaemia cells under a microscope. A leukaemia cell count of less than 5 in a hundred cells has been considered a remission. However it does not need an expert to realise that even low levels of leukaemia remaining are still significant hence the use of Minimal residual disease (MRD) testing which uses highly sensitive molecular testing (usually PCR) to evaluate low levels of leukaemia cells usually down to a level of $1x10^{-5}$. I would consider a negative MRD test as a clinically significant treatment response.
9. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes definitely.
What is the expected place of the technology in current practice?	

10. How is the condition currently treated in the NHS?	Standard chemotherapy protocols based on the UKALL14 trial +/- stem cell transplantation.
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	No UK National Guidelines but some people use the American NCCN Guidelines
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	Across the UK treatment pathways and protocols are well defined
• What impact would the technology have on the current pathway of care?	This intervention would be used for patients who are found to be MRD positive after 2 cycles of standard induction chemotherapy. Blinatumomab has not been widely used outside of allogeneic transplant centres and so would likely result in an increase in referrals to these centres for this treatment.
11. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Yes.

• How does healthcare resource use differ between the technology and current care?	This technology requires the use of special ambulatory pumps (so the patients can receive part of the treatment out of hospital), these are generally only used in larger centres with ambulatory care teams.
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	This should definitely be used in the allogeneic transplant centres who have both the infrastructure and experience.
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	any new centres would need ambulatory pumps, training in their use and training in the administration of/care of patients receiving Blinatumomab
12. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes.
• Do you expect the technology to increase length of life more than current care?	Yes

• Do you expect the technology to increase health-related quality of life more than current care?	Yes
13. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	No
The use of the technology	
14. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors	This is definitely more difficult to use than current treatment (as above)

affecting patient acceptability or ease of use or additional tests or monitoring needed.)	
15. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	This should be available for patients who remain MRD positive (at any level) following 2 cycles of standard induction chemotherapy. Up to 2 cycles of treatment should be given followed by a reassessment bone marrow MRD test (this is an additional test)
16. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	No
17. Do you consider the technology to be innovative in its potential to make a significant and substantial	Yes

impact on health-related benefits and how might it improve the way that current need is met?	
• Is the technology a 'step-change' in the management of the condition?	Yes. Compared to conventional chemotherapy this treatment is very well tolerated.
• Does the use of the technology address any particular unmet need of the patient population?	Yes. Patients who are MRD positive after 2 cycles of standard induction chemotherapy have a poor outlook.
18. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	Blinatumomab is generally very well tolerated compared to conventional chemotherapy and patients QOL generally improves whilst they are receiving it.
Sources of evidence	

19. Do the clinical trials on the technology reflect current UK clinical practice?	Yes
• If not, how could the results be extrapolated to the UK setting?	N/A
• What, in your view, are the most important outcomes, and were they measured in the trials?	The most important outcome is to eliminate low levels of leukaemia cells i.e render the patients bone marrow test 'MRD negative'
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	Yes. Failing to achieve MRD negativity is associated with a poor outcome.
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	No
20. Are you aware of any relevant evidence that might	No

not be found by a systematic review of the trial evidence?	
22. How do data on real-world experience compare with the trial data?	I am not aware of any published data comparing real world experience in adults but personal experience suggests that outcomes are similar.
Equality	
23a. Are there any potential equality issues that should be taken into account when considering this treatment?	No
23b. Consider whether these issues are different from issues with current care and why.	
Key messages	

25. In up to 5 bullet points, please summarise the key messages of your statement.

- Blinatumomab is a highly effective treatment with elimination of MRD in approx. 80% of patients in CR1 (BLAST study)
- Blinatumomab is well tolerated compared to conventional cytotoxic chemotherapy.
- A Phase 3 study showed increase QOL measures with Blinatumomab compared to reduced QOL measures in patients receiving chemotherapy.
-
-

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

Blinatumomab for acute lymphoblastic leukaemia for people with minimal residual disease activity in remission: A Single Technology Appraisal

Produced by	School of Health and Related Research (ScHARR), The University of Sheffield
Authors	Jean Hamilton, Research Fellow, ScHARR, University of Sheffield, UK Daniel Pollard, Research Associate, ScHARR, University of Sheffield, UK Emma Simpson, Senior Research Fellow, ScHARR, University of Sheffield, UK Paul Tappenden, Reader in Health Economic Modelling, ScHARR, University of Sheffield, Sheffield, UK Ruth Wong, Information Specialist, ScHARR, University of Sheffield, Sheffield, UK Joanna Leaviss, Research Fellow, ScHARR, University of Sheffield, UK Clare Rowntree, Consultant Haematologist, University Hospital of Wales, Cardiff, UK Tobias Menne, Consultant Haematologist, Freeman Hospital, Newcastle, UK
Correspondence Author	Jean Hamilton, ScHARR, University of Sheffield, UK

Date completed	30th January 2018

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 16/112/08.

Declared competing interests of the authors

Tobias Menne is a member of three independent data monitoring committees for studies sponsored by Amgen, but for different interventions than that considered in the current appraisal. None of the other authors have any conflicts of interest to declare.

Acknowledgements

We would like to thank Sarah Davis, Senior Lecturer, ScHARR, for providing comments on the draft report and Gill Rooney, Programme Manager, ScHARR, for providing administrative support and in preparing and formatting the report.

Rider on responsibility for report

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

This report should be referenced as follows:

Hamilton J, Pollard D, Simpson EL, Tappenden P, Wong R, Leaviss J, Rowntree C, Menne T. Blinatumomab for acute lymphoblastic leukaemia for people with minimal residual disease activity in remission: A Single Technology Appraisal. Final report to the National Institute for Health and Care Excellence. School of Health and Related Research (ScHARR), University of Sheffield, 2018.

Contributions of authors

Emma Simpson and Joanna Leaviss summarised and critiqued the clinical effectiveness evidence reported within the company's submission. Jean Hamilton critiqued the statistical analyses undertaken by the company. Paul Tappenden and Daniel Pollard critiqued the health economic analysis submitted by the company. Ruth Wong critiqued the company's search strategy. Clare Rowntree and Tobias Menne provided clinical advice to the ERG throughout the project. All authors were involved in drafting and commenting on the final report.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF BOXES

Abbreviations

AE	Adverse event
ANC	Absolute neutrophil count
AIC	Akaike information criterion
ALL	Acute lymphoblastic leukaemia
AML	Acute myeloid leukaemia
ATE	Average treatment effect
ATT	Average treatment effect on the treated
BCP	B-cell precursor
BIC	Bayesian Information Criterion
BNF	British National Formulary
BSA	Body surface area
CEAC	Cost-effectiveness acceptability curves
CI	Confidence interval
CMU	Commercial Medicines Unit
CMV	Cytomegalovirus
CNS	Central nervous system
CR	Complete remission
CRD	Centre for Reviews and Dissemination
CS	Company's submission
CSF	Cerebrospinal fluid
DCAS	Direct comparison analysis set
DES	Discrete event simulation
DFS	Disease-free survival
DSA	Deterministic sensitivity analyses
ECOG	Eastern Cooperative Oncology Group
EFS	Event-free survival
EMBASE	Excerpta Medica dataBASE
eMIT	Electronic Market Information Tool
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire C30
EPAR	European Public Assessment Report
EQ-5D	Euroqol 5-Dimensions
ERG	Evidence Review Group
FCE	Finished consultant episode
GEE	Generalised estimating equation
GLM	Generalised linear model
GMALL	German Multicenter Acute Lymphoblastic Leukaemia Study Group
HB	Haemoglobin
HR	Hazard ratio
HRQoL	Health-related quality of life

HSCT	Haematopoietic stem cell transplantation
HST	Highly Specialised Technologies
ICER	Incremental cost-effectiveness ratio
IPD	Individual patient-level data
IPTW	Inverse probability of treatment weighting
IV	Intravenous
Kg	Kilogram
MEDLINE	Medical Literature Analysis and Retrieval System Online
Mg	Milligram
MRD	Minimal residual disease
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIH	National Institutes of Health
NSAIDS	Non-steroidal anti-inflammatory drugs
OS	Overall survival
PAS	Primary analysis set
PCR	Polymerase chain reaction
Ph	Philadelphia chromosome
PRO	Patient-reported outcome
PRS	Post-relapse survival
PSA	Probabilistic sensitivity analyses
PSS	Personal Social Services
QALY	Quality-adjusted life year
QuEENS	Quality of Effectiveness Estimates from Non-randomised Studies
RCS	Restricted cubic spline
RCT	Randomised controlled trial
RFS	Relapse-free survival
ROBINS	Risk Of Bias In Non-randomised Studies
R/R	Relapsed/refractory
RT-qPCR	Real-time quantitative polymerase chain reaction
SAE	Serious adverse event
sATT	Stabilised average treatment effect on the treated
SC	Standard care
SD	Standard deviation
SE	Standard error
TA	Technology appraisal
TCR	T-cell receptor
TKI	Tyrosine kinase inhibitor
WBC	White blood cells
WTP	Willingness-to-pay

1. SUMMARY

1.1 Critique of the decision problem in the company's submission

The company's submission (CS) assesses the clinical effectiveness and cost-effectiveness of blinatumomab (Blincyto®), within its anticipated licensed indication for the treatment of adult patients with minimal residual disease-positive B-cell precursor acute lymphoblastic leukaemia (MRD+ BCP-ALL) whilst in remission. The company's description of ALL and its management is broadly appropriate. The decision problem addressed by the CS is partly in line with the final scope issued by the National Institute for Health and Care Excellence (NICE). The indirect comparison and health economic analysis presented within the CS compare blinatumomab with standard care chemotherapy within a population of adult patients with Philadelphia chromosome-negative (Ph-) disease with first complete haematological remission (CR1); this is narrower than the population defined by the anticipated license indication for blinatumomab. As such, the company's indirect comparison and health economic analysis exclude two groups of patients who were enrolled into the BLAST study: (i) patients who are in second or subsequent haematological remission (CR2+), and (ii) patients with Ph+ ALL (any CR). Despite this absence of evidence, the CS argues that due to the substantial unmet need across all subgroups, blinatumomab should be considered for use within its full anticipated marketing authorisation. However, the company further suggests that blinatumomab should be used early in the treatment pathway, with initiation after front-line chemotherapy (after two induction cycles) for those patients with persistent MRD at this stage. The CS also excludes the comparator of "monitor for relapse" based on the argument that it is highly unlikely that MRD+ patients who are at high risk of relapse would not receive active treatment. However, clinical advisors to the Evidence Review Group (ERG) noted that due to its favourable toxicity profile, blinatumomab may be a potential treatment option for patients who are unable to undergo haematopoietic stem cell transplantation (HSCT) or to tolerate chemotherapy; the ERG considers that a further comparison of blinatumomab versus monitoring within this subgroup should have been explored.

1.2 Summary of clinical effectiveness evidence submitted by the company

The key clinical effectiveness evidence for blinatumomab was based on two single-arm open-label studies; BLAST (n=116) and the pilot study MT103-202 (n=20). From the 116 patients in BLAST, median overall survival (OS) was [redacted], with an 18-month OS probability of [redacted]. From 110 patients providing relapse-free survival (RFS) data from BLAST, median RFS was [redacted], with an 18-month RFS probability of [redacted]. Based on the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire C30 (EORTC QLQ-C30), patients reported [redacted]. By the end of the core study,

[redacted]

HRQoL as measured by the Euroqol 5-Dimensions (EQ-5D) questionnaire did not change significantly from baseline to the end of the core study. [redacted] experienced at least one treatment-emergent AE.

Comparator data relating to standard care chemotherapy were provided from one historical control study, Study 20120148 (n=287); this study was based on data obtained from existing clinical databases.

Owing to the lack of randomised data to inform the comparative effectiveness of blinatumomab versus standard care chemotherapy, treatment effects were estimated using non-randomised data from BLAST and the historical control study. Due to differences between the populations of BLAST and the historical control study, comparative analyses were undertaken using subsets of the original study populations which were restricted to patients with Ph- disease in CR1 only: the BLAST primary analysis set (PAS, [redacted]) and the historical control direct comparison analysis set (DCAS, [redacted]). A propensity score model was constructed and used to generate weights which were applied to the historical control DCAS, with the aim of approximating the response to standard care chemotherapy that would be expected in a population with the same characteristics as the BLAST PAS. The resulting average treatment effect on the treated (ATT) estimates are applicable to Ph- and CR1 individuals only. This analysis suggested a hazard ratio (HR) [redacted].

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

Despite limitations in the company's search strategy, the ERG considers it unlikely that any relevant studies of blinatumomab in adult BCP-ALL patients with MRD positivity after treatment have been missed by the company's searches. The eligibility criteria applied in the selection of evidence for the clinical effectiveness review were considered by the ERG to be reasonable and consistent with the decision problem outlined in the final NICE scope, with the exception that the comparator "monitor for relapse" was not included. The ERG's clinical advisors noted that some older and less fit patients may not be able to receive HSCT or to tolerate chemotherapy, but may be able to tolerate blinatumomab, and so this comparator is potentially relevant for a subgroup of MRD+ BCP-ALL patients. It is unclear whether potentially relevant comparator data exist for this subgroup (for example, from registry sources).

The main evidence in the CS was from the single-arm BLAST study. Whilst BLAST was generally well reported and conducted, single-arm studies are associated with an array of potential biases

including a high risk of selection bias (due to the absence of randomisation), performance bias and detection bias (due to the absence of blinding).

The ERG considers that the propensity score methods used by the company to inform comparative effectiveness estimates were appropriate. However, the estimation of treatment effects based on non-randomised data is still subject to inherent limitations, namely that it is not possible to account for unobserved confounders. It was unclear whether the uncertainty associated with the propensity score weights was accounted for when estimating the treatment effects. The ERG therefore considers that the reported treatment effects are likely to underestimate the associated uncertainty and should be interpreted with caution. There was also a lack of clarity and consistency in the weighted analyses presented within the CS, as results using stabilised ATT (sATT) weights were presented in the clinical effectiveness section, and standard (non-stabilised) weights were used to inform the health economic model.

The key uncertainties in the clinical evidence relate to the comparative efficacy and the generalisability of the available evidence to the full population outlined in the scope.

1.4 Summary of cost effectiveness submitted evidence by the company

The company's *de novo* partitioned survival model assesses the cost-effectiveness of blinatumomab versus chemotherapy (based on the UK ALL14 maintenance regimen) in patients with Ph- MRD+ BCP-ALL in CR1. Incremental health gains, costs and cost-effectiveness of blinatumomab are evaluated over a 50-year time horizon from the perspective of the National Health Service (NHS) and Personal Social Services (PSS). The company's model is comprised of a main structure which reflects RFS and OS outcomes, together with two linked sub-models which are intended to estimate additional costs and HRQoL decrements associated with HSCT received before and/or after relapse. The main model structure includes three health states: (1) relapse-free; (2) post-relapse and (3) dead. The survival models were generated from analyses of time-to-event data (RFS and OS) from the company's propensity score analysis of the BLAST PAS and the historical control study DCAS using ATT weights. RFS is modelled using an unrestricted Gompertz distribution (equivalent to fitting separate models to both groups), whilst OS is modelled using a log normal mixture cure model (whereby the parameters of the log normal distribution are the same for both groups, but the cure fraction is allowed to differ between the groups). HRQoL is assumed to be principally determined by relapse status, time spent in the relapse-free state and treatment received; utility estimates were derived from a generalised linear model/generalised estimating equation (GLM/GEE) model fitted to EQ-5D data collected in BLAST, a further propensity matching analysis of the BLAST and TOWER blinatumomab studies, as well as other literature and assumptions. Resource use estimates and costs were based on data collected in BLAST, the UK ALL14 treatment protocol, routine cost sources, clinical opinion and other literature.

Based on the probabilistic version of the model (assuming the unrestricted Gompertz function for RFS and the log normal mixture cure model for OS), blinatumomab is expected to generate an additional 2.85 quality-adjusted life years (QALYs) at an additional cost of £84,456 compared with standard care: the corresponding incremental cost-effectiveness ratio (ICER) for blinatumomab versus standard care is £29,673 per QALY gained. The deterministic version of the company's model produces a similar ICER of £28,524 per QALY gained for blinatumomab versus standard care. Assuming a willingness-to-pay (WTP) threshold (λ) of £20,000 per QALY gained, the company's model suggests that the probability that blinatumomab produces more net benefit than standard care is 0.10; assuming a WTP threshold of £30,000 per QALY gained, the company's model suggests that the probability that blinatumomab produces more net benefit than standard care is 0.53. Following the clarification process, the company submitted a revised model which addressed some of the minor concerns initially raised by the ERG; the probabilistic version of the company's updated model suggests that the ICER for blinatumomab versus standard care is £28,655 per QALY gained.

1.5 Summary of the ERG's critique of cost effectiveness evidence submitted

The ERG critically appraised the company's economic analysis and double-programmed the deterministic version of the company's model. The ERG's critical appraisal identified several issues relating to the company's economic analysis and the evidence used to inform it. These include: (i) the exclusion of relevant patient subgroups from the model; (ii) the exclusion of the "monitor for relapse" comparator from the analysis; (iii) use of a model structure which is inappropriate for tracking HSCT; (iv) the absence of RCT evidence for blinatumomab versus standard care; (v) concerns regarding the company's approach to RFS/OS model selection; (vi) concerns regarding the robustness of the company's alternative base case (blinatumomab used on relapse for the standard care group); (vii) the questionable reliability of the company's HRQoL estimates; (viii) uncertainty surrounding the proportion of RFS events that are deaths; (ix) the inclusion of an unrealistic treatment pathway and (x) limited sensitivity analysis around alternative parametric functions.

1.6 ERG commentary on the robustness of evidence submitted by the company

1.6.1 Strengths

The ERG considers it unlikely that any relevant studies of blinatumomab have been missed the company's searches. The BLAST study was a well conducted single-arm study which reported on the full range of outcomes listed in the NICE scope (although comparative analyses from the company's propensity score model were restricted to RFS and OS outcomes only).

The CS details extensive efforts taken to verify the correct implementation of the health economic model and to ensure the accuracy of the parameter inputs against the source material from which these were derived. The company's model was found to include only minor errors.

1.6.2 Weaknesses and areas of uncertainty

The key weaknesses in the evidence base relate to the lack of randomised evidence to inform comparative effectiveness and the limited generalisability of the available evidence to the full population defined by the NICE scope and the anticipated license authorisation. The ERG considers the following to represent the key uncertainties within the clinical and economic evidence base for blinatumomab:

- The absence of comparative clinical and economic evidence for blinatumomab versus standard care chemotherapy within subgroups of the BLAST study which were excluded from the comparative analysis (patients with Ph+ MRD+ BCP-ALL and patients with Ph- MRD+ BCP-ALL with CR2+).
- The absence of clinical data and economic comparisons of blinatumomab versus monitoring for patients who are unable to undergo HSCT or to tolerate chemotherapy.
- The necessary reliance on adjusted historical control evidence, due to the absence of RCT evidence for blinatumomab versus standard care, and the potential for unobserved confounders.
- The long-term extrapolation of RFS and OS outcomes, including the timing of cure.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG undertook eight sets of exploratory analyses using the deterministic version of the company's updated model. Notwithstanding uncertainty relating to the choice of parametric RFS and OS functions, the ERG's preferred model includes the correction of seven minor programming errors and the inclusion of a fixed 5-year cure point. The ERG-preferred model produces a deterministic ICER for blinatumomab versus standard care of £30,227 per QALY gained. The ERG also undertook a number of further analyses to explore the impact of alternative parametric models and alternative parameter values on the results of the ERG-preferred model. These analyses indicate that the costs of standard care chemotherapy, the post-HSCT survival probabilities and the utility value for the post-relapse state have only a minor impact on the ICER for blinatumomab versus standard care. Conversely, the cure fraction and the choice of parametric OS distribution have a significant impact on the model results. Within the ERG's exploratory analysis of alternative RFS and OS models, the ICER for blinatumomab versus standard care ranges from £25,783 per QALY gained (Weibull non-mixture cure model, unrestricted) to £63,265 per QALY gained (Weibull model, unrestricted). Across the full range of models considered, only the Weibull non-mixture cure model (unrestricted) and the Weibull mixture cure model (unrestricted) produce results in which the full range of deterministic ICERs are below £30,000 per QALY gained (irrespective of RFS model assumed). The clinical advisors' three preferred

OS models (the generalised gamma [unrestricted], the restricted cubic spline (RCS) Weibull [unrestricted] and the Weibull mixture cure [unrestricted]) result in deterministic ICERs in the range £25,810 per QALY gained to £34,904 per QALY gained.

On the basis of the results of the 35 parametric OS models considered within the ERG's exploratory analyses, the ERG does not believe that blinatumomab meets NICE's criteria for life-extending treatments given at the end of life.

2. BACKGROUND

This report provides a review of the evidence submitted by Amgen in support of blinatumomab for acute lymphoblastic leukaemia (ALL) for people with minimal residual disease (MRD) activity in remission. It considers both the original company submission (CS)[1] received on 8th November 2017 and a subsequent response to clarification questions supplied by Amgen on 13th December 2017.

2.1 Critique of company's description of the underlying health problem

The CS[1] (pages 19-31) provides a reasonable description of the underlying health problem; this is summarised briefly below.

ALL is a rare and rapidly progressing form of leukaemia characterised by the excess production of immature lymphocyte precursor cells, called lymphoblasts or blasts cells, in the bone marrow. Lymphocytes are white blood cells that are vital for the body's immune system. Eventually, this affects the production of normal blood cells which leads to a reduction in the numbers of red cells, white cells and platelets in the blood.[1] ALL represents about 20% of all leukaemias in adults.[2, 3]

There are a number of sub-classifications of ALL, with the majority (approximately 76%) of adult cases being B-cell lineage (based on a weighted average of five estimates synthesised by the company[4, 5 6 7 8]). Of these, approximately 93% are B-cell precursor (BCP) ALL (based on a weighted average of two studies[9 10]). Therefore, BCP-ALL constitutes approximately 71% of the adult ALL population, which is expected to equate to around 236 patients in England and Wales.[1] Approximately 25%[3, 9] of adults with ALL (across all sub-classifications, not specifically BCP) have an acquired chromosomal abnormality, known as Philadelphia chromosome-positive (Ph+) disease, which is caused by reciprocal translocations between chromosomes 9 and 22. These translocations result in a BCR-ABL fusion gene that encodes an active tyrosine kinase protein which causes uncontrolled cell proliferation. The presence of the Ph chromosome in adults increases with age[2, 3, 11] and Ph+ ALL individuals typically have a worse prognosis than those without the abnormality.[12]

Many of the patients who achieve the criteria for haematological complete remission (CR) will experience a recurrence of disease; this is thought to result from residual leukaemia cells that remain.[1] MRD describes residual ALL in patients in CR that is detectable only by molecular means.[13] Patients are considered to have clinically significant MRD, and are described as being MRD+,[3, 14] if their MRD level is greater than 1 x 10^{-4}, although clinical studies have assessed MRD positivity using various thresholds. The company estimates that 36% of all BCP-ALL patients in CR exhibit MRD+, based on a weighted analysis of Ph- patients in three studies;[4, 15, 16] this implies an estimated 85 cases of MRD+

BCP-ALL in England and Wales.[1] The company's clarification response[17] (question A2) estimates that approximately 15 of these patients will be Ph+.

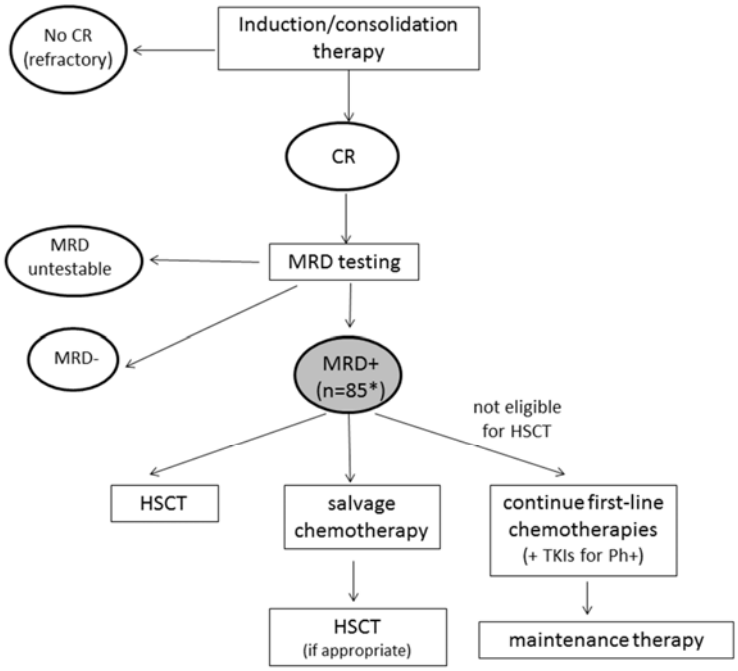
The prognosis for patients with BCP-ALL is dependent on a number of factors. Well-established positive prognostic factors include: younger age; shorter time to CR; longer duration of CR; absence of poor risk cytogenetics such as Ph+, and lower white blood cell counts.[3, 5, 7] In addition, MRD positivity is a major and well established risk factor.[13] In a large German Multicentre Acute Lymphoblastic Leukaemia (GMALL) study of Ph- ALL,[4] the probability of overall survival (OS) at 5 years was 42% for MRD+ compared with 80% for MRD- patients (MRD assessed at week 16 after consolidation therapy). In a meta-analysis by Berry *et al*,[18] poorer outcomes for MRD+ patients compared with MRD- patients were observed. Although OS estimates were not reported specifically for the MRD+ BCP-ALL subgroup, the persistence of MRD was shown to be a strong predictive factor for relapse and OS, irrespective of ALL cell phenotype (B-cell or T-cell), Ph chromosome subgroup and MRD detection method, cut-off, or timing of assessment.[18]

2.2 Critique of company's overview of current service provision

In general, the CS[1] (pages 26-30) provides a reasonable overview of current service provision for people with MRD+ BCP-ALL, although the submission is not always clear where information relates to specific sub-populations of ALL patients. The company's description of the treatment pathway is briefly summarised in this section, and is supplemented with information provided by clinical advisors to the Evidence Review Group (ERG).

The management of people with MRD+ BCP-ALL is complex and there is currently no guidance published by the National Institute for Health and Care Excellence (NICE) for the treatment of adults with MRD+ BCP-ALL in England. The treatment of ALL in the UK is generally based on the UKALL14 protocol.[19] In general, the treatment approach varies according to age, general fitness and health at diagnosis and the results of cytogenetic testing.[2] The aim of treatment is to achieve cure (defined as sustained MRD negativity) and maintained haematological CR (defined as a bone marrow blast level of <5%[13]). According to clinical advice received by the ERG, patients who do not experience relapse within 5 years of diagnosis are generally considered to be cured. Most long-term survivors achieve cure by undergoing haematopoietic stem cell transplantation (HSCT), although this is not always required for standard-risk patients, and some high-risk patients may not be suitable candidates for HSCT for various reasons, for example, due to older age, medical comorbidities, or the lack of a suitable donor.[13] Figure 1 presents an overview of the treatment pathway for people with MRD+ BCP-ALL.

Figure 1: Overview of the treatment pathway for MRD+ BCP-ALL



** Estimated number of patients in England and Wales from CS,[1] The grey box indicates the relevant population for this appraisal*

In clinical practice, most treatment plans have three phases: (i) induction (with or without intensification); (ii) consolidation, and (iii) maintenance (in adults, later stages of treatment may be replaced by allogeneic transplantation). The aim of induction therapy is to achieve full remission quickly. Patients are treated with established standard chemotherapy combinations (including tyrosine-kinase inhibitor [TKI] therapy for Ph+ patients only). Once in remission, patients may proceed to HSCT, with or without intensification, if considered high-risk for relapse (e.g. MRD+, poor risk cytogenetics, age over 40 years) assuming they are clinically eligible, willing to undergo HSCT and have a suitable donor. Currently, adult patients with a sibling donor would also undergo HSCT in first remission. Consolidation therapy followed by maintenance therapy is given to patients who are not eligible for HSCT or who have standard-risk disease and no sibling donor. Ph+ patients additionally receive daily imatinib (a TKI therapy) throughout induction and intensification. As patients with Ph+ disease are deemed high-risk, they would usually have an HSCT instead of ongoing consolidation and maintenance chemotherapy unless they are not considered fit enough for transplant or do not have a suitable donor.

MRD testing is widely implemented in the UK and is recommended as standard care in the patient management process for ALL.[13] However, global consensus has not yet been reached on when to test

for MRD. Brüggemann *et al*[20] determined that the timing of MRD status influences outcomes, with patients achieving MRD negativity during induction experiencing improved relapse-free survival (RFS) and OS compared with patients achieving MRD negativity after induction. A survey of 20 UK physicians undertaken by the company[21] suggested an apparent consensus on MRD testing patterns in the UK. Based on the survey data, an initial prognostic MRD test was commonly conducted 4-8 weeks after the start of induction therapy. Once a patient has achieved MRD negativity, they do not have further testing if they remain on chemotherapy only. Patients undergoing transplantation will have an average of 4 post-CR MRD tests, at roughly 3 month intervals, over the subsequent 12 months post-transplant (irrespective of MRD status pre transplant). The rare patients that continue chemotherapy despite being MRD+ (due to patient choice, fitness or lack of donor) would not receive further routine MRD testing due to the lack of current options for curative treatment post-relapse.

The company suggests that blinatumomab should be used early in the treatment pathway, with initiation after front-line chemotherapy (after two induction cycles) for those patients with persistent MRD at this stage. According to the CS,[1] blinatumomab is expected to displace continued chemotherapy and/or be used prior to HSCT. Blinatumomab is not intended to displace HSCT, rather it is likely to be used prior to HSCT in patients who are eligible to undergo transplant, with the aim of increasing the likelihood of a positive outcome, or to delay the need for HSCT.[22] Despite this, the company suggests that by achieving and sustaining MRD negativity over time, blinatumomab may conceivably delay transplant indefinitely (the ERG notes that this argument suggests that blinatumomab would displace HSCT, at least in some patients).

3. CRITIQUE OF COMPANY'S DEFINITION OF THE DECISION PROBLEM

This chapter presents a summary and critique of the decision problem addressed by the CS.[1] A summary of the decision problem as outlined in the final NICE scope[22] and addressed in the CS[1] is presented in Table 1.

Table 1: Company's statement of the decision problem (reproduced from CS Table 1)

	Final scope issued by NICE[22]	Decision problem addressed in the CS[1]	Company's rationale if different from the final NICE scope
Population	People with BCP-ALL who have MRD activity while in haematological remission	Adults with MRD+ B-precursor ALL. Clinical evidence for blinatumomab is aligned with the proposed licensed indication; however, comparative evidence from a historical comparator study is limited to patients with Ph-negative B-precursor ALL who are in first complete haematological remission. Therefore, the economic analysis presented in this submission focused on this patient subgroup. Although the cost-effectiveness evidence does not consider the Ph+ population or later remission states, due to the substantial unmet need across all sub-populations blinatumomab should be considered for use in alignment with its full anticipated marketing authorisation.	Blinatumomab is not expected to have a marketing authorisation for use in paediatric patients in this indication.
Intervention	Blinatumomab	As per final scope	N/a
Comparator(s)	• Retreatment with combination chemotherapy • Monitor for relapse	• Retreatment with combination chemotherapy	Based on expert clinical opinion it is highly unlikely that MRD+ patients who have a high-risk of relapse would solely be monitored for relapse without any treatment. Therefore, in the economic evaluation monitoring for relapse is not considered a comparator in its own right – instead, it is captured alongside ongoing chemotherapy regimens.
Outcomes	The outcome measures to be considered include: • Overall survival • Disease-free survival • Relapse-free survival • MRD response • Rate of stem cell transplant • Adverse effects of treatment • HRQoL	As per final scope	N/a

	Final scope issued by NICE[22]	Decision problem addressed in the CS[1]	Company's rationale if different from the final NICE scope
Special considerations including issues related to equity or equality	If appropriate, the appraisal should include the costs associated with diagnostic testing for these cells in people with ALL, while in remission, who would not otherwise have been tested. A sensitivity analysis should be provided without the cost of the diagnostic test. Guidance will only be issued in accordance with the marketing authorisation. Where the wording of the therapeutic indication does not include specific treatment combinations, guidance will be issued only in the context of the evidence that has underpinned the marketing authorisation granted by the regulator.	As per final scope	MRD status testing is already routine clinical practice in the diagnostic work-up and monitoring of BCP-ALL,[13, 23] and is recognised as an important marker for informing treatment decisions and prognosis. No additional tests or investigations are required for treatment with blinatumomab.

N/a - Not applicable

3.1 Population

The population defined in the final NICE scope[22] relates to people with BCP-ALL who have MRD activity while in remission. Blinatumomab does not currently have a marketing authorisation for this indication. According to the draft SmPC submitted to NICE by the company,[24] the anticipated wording of the marketing authorisation is as follows: *"BLINCYTO [blinatumomab] is indicated for the treatment of adults with minimal residual disease (MRD) positive B precursor ALL."* This population is in line with the BLAST study,[1] but relates only to adult ALL patients. The ERG notes that the indirect comparison and the health economic analysis presented within the CS (see Sections 4.4 and 5.2, respectively) relate to a narrower population of adult patients with Ph- disease with first complete haematological remission (CR1). Consequently, the indirect comparison and health economic analysis exclude two groups of patients who were enrolled into BLAST and who are included in the anticipated marketing authorisation: (i) patients who are in second or subsequent haematological remission (CR2+), and (ii) patients with Ph+ ALL (any CR). Despite this absence of evidence, the CS argues that due to the substantial unmet need across all subgroups, blinatumomab should be considered for use within its full anticipated marketing authorisation. In addition, clinical advisors noted that due to its toxicity profile, blinatumomab represents a potential treatment option for patients who are unable to undergo HSCT or to tolerate chemotherapy; it is unlikely that this subgroup is reflected within the population of patients enrolled into the BLAST study. These issues are discussed further in Section 5.3.

The ERG notes that the CS does not include any clinical or economic evidence relating to paediatric patients; the draft SmPC for blinatumomab[24] notes that the safety and efficacy of blinatumomab in paediatric patients have not yet been established. The draft SmPC also states that there is limited experience with blinatumomab in patients ≥75 years of age, and that the safety and efficacy of blinatumomab have not been studied in patients with severe renal impairment or in patients with severe hepatic impairment.[24]

3.2 Intervention

The intervention under appraisal is blinatumomab (Blincyto®). Blinatumomab is a bispecific T-cell engager antibody construct that binds specifically to CD19 expressed on the surface of cells of B-lineage origin and CD3 expressed on the surface of T-cells.[24] Blinatumomab currently holds an EU marketing authorisation for the treatment of adults with Ph- relapsed or refractory (R/R) BCP-ALL. The CS[1] highlights that blinatumomab is the first and only drug indicated specifically for MRD+ BCP-ALL patients in haematological CR.

Blinatumomab is available as a single vial containing 38.5µg of blinatumomab solution. The current list price for a single vial of blinatumomab is £2,017.[25] A simple discount Patient Access Scheme has

been approved by the Department of Health; including the discount, the price of blinatumomab is [redacted] per vial.[1]

Within its MRD+ BCP-ALL indication, patients may receive one cycle of induction treatment followed by up to three additional cycles of blinatumomab consolidation treatment. A single cycle of treatment is comprised of 28 days of continuous intravenous (IV) infusion followed by a 14-day treatment-free interval.[24] The draft SmPC states that when considering the use of blinatumomab as a treatment for MRD+ BCP-ALL, detectable MRD (defined as molecular relapse or molecular failure) should be confirmed in a validated assay with minimum sensitivity of 10^{-4}. Clinical testing of MRD, regardless of the choice of technique, should be performed by a qualified laboratory familiar with the technique.[24]

The draft SmPC[24] states that the decision to discontinue blinatumomab temporarily or permanently, as appropriate, should be made in the case of the following severe (Grade 3) or life-threatening (Grade 4) toxicities: cytokine release syndrome; tumour lysis syndrome; neurological toxicity; elevated liver enzymes, and any other clinically relevant toxicities (as determined by the treating physician).

The draft SmPC[24] lists the following special warnings and precautions for use: neurologic events; infections; cytokine release syndrome and infusion reactions; tumour lysis syndrome; neutropenia and febrile neutropenia; elevated liver enzymes; pancreatitis; leukoencephalopathy including progressive multifocal leukoencephalopathy; immunisations; contraception; medication errors and excipients with known effect.

Contraindications to blinatumomab include hypersensitivity to the active substance or to any of the excipients listed in the SmPC and breast-feeding.[24]

3.3 Comparators

The final NICE scope[22] defines two relevant comparators: (i) retreatment with combination chemotherapy, and (ii) monitor for relapse.

The company's review of clinical effectiveness (see CS,[1] Section B2) did not identify any studies which included head-to-head comparisons of blinatumomab versus either of the comparators listed in the final NICE scope.[22] As a consequence, the company's systematic review focusses on a single historical control comparator study that included adult Ph- BCP-ALL patients who have received country-specific standard care treatments (according to the locations in which the study was conducted), achieved a haematological CR, and subsequently had persistent or relapsed MRD. The range of chemotherapy regimens received by patients within the historical control study is not reported, however, the CS refers to *"standardised treatment protocols developed as part of the European Working Group for Acute*

Lymphocytic Leukaemia (EWALL) collaboration."[1] Clinical advice received by the ERG suggests that this should ensure that patients are treated to a similar standard across countries. The company's model uses propensity score methods with average treatment effect on the treated (ATT) weights to adjust the observed data for the standard care group to reflect the characteristics of the blinatumomab study (BLAST). The model assumes that standard care chemotherapy is given according to the UKALL14 trial maintenance regimen;[19] this is comprised of: (i) vincristine (IV, 1.4mg/m^2 once every 13 weeks); (ii) methotrexate (intrathecal, 12.5mg once every 13 weeks); (iii) prednisolone (oral, 60mg/m^2 5 times every 13 weeks); (iv) mercaptopurine (oral, 75mg/m^2 daily), and (v) methotrexate (oral, 20mg/m^2 weekly). This regimen is used only to estimate the costs of chemotherapy; downstream interventions within both treatment groups include allogeneic HSCT (given pre- and/or post-relapse) and salvage chemotherapy using the FLAG-IDA regimen.

As shown in Table 1, the CS does not consider evidence relating to the "monitor for relapse" comparator; this exclusion is based on the argument that it is highly unlikely that MRD+ patients who have a high risk of relapse would solely be monitored for relapse without any treatment. As noted in Section 3.1, the clinical advisors to the ERG suggested that some older and less fit patients will not be able to undergo HSCT or to tolerate chemotherapy, but may be offered blinatumomab for the treatment of persistent MRD positivity. Therefore, monitoring for relapse is a relevant comparator within this patient subgroup and should have been considered in the CS.

3.4 Outcomes

The final NICE scope[22] lists the following outcomes:

- Overall survival (OS)
- Disease-free survival (DFS)
- Relapse-free survival (RFS)
- Minimal residual disease (MRD) response
- Rate of stem cell transplant
- Adverse effects of treatment
- Health-related quality of life (HRQoL)

The CS[1] reports on all of these outcomes for patients receiving blinatumomab within the BLAST study. The reporting of outcomes for the historical control study is restricted to RFS and OS (see CS,[1] Section B.2.9.4). The company's health economic model is based on data from BLAST and the historical control study relating to RFS, OS, HSCT rates and HRQoL.

3.5 Economic analysis

The CS reports the methods and results of a *de novo* model-based health economic analysis to assess the incremental cost-effectiveness of blinatumomab versus standard care chemotherapy for the treatment of adults with MRD+ B-precursor Ph- ALL in CR1. The company's health economic analysis is detailed and critiqued in Chapter 5.

3.6 Subgroups

The final NICE scope[22] does not specify any subgroups of patients with MRD+ BCP-ALL. The company's indirect comparison and health economic analysis are restricted to patients with MRD+ BCP-ALL who are Ph- and in CR1. The company's clinical effectiveness review includes an analysis of RFS and OS outcomes for BLAST patients in CR2 according to MRD response; however, no comparative analyses are presented against other standard care therapies.

3.7 Special considerations

The CS[1] states that there are no equality issues relating to the use of blinatumomab for the treatment of adult MRD+ BCP-ALL patients in haematological CR.

The CS states that blinatumomab is indicated for a rare condition which affects only a very small number of patients (85 patients per year). According to the CS, these patients have a significant unmet medical need and they may gain substantially from access to blinatumomab. The CS goes on to argue that blinatumomab meets many of the criteria for appraisal under the Highly Specialised Technologies (HST) framework and as such, blinatumomab should be evaluated taking into account a wider range of criteria relating to benefits and costs. As blinatumomab has been referred for appraisal under the Technology Appraisal (TA) programme, this issue is not discussed further within this ERG report.

The CS also claims that on the basis of median OS gains derived from the ATT-weighted propensity score analyses, blinatumomab meets NICE's criteria for life-extending treatments given at the end of life.[26] Undiscounted mean OS estimates for blinatumomab and standard care are not used to support this argument, but can be generated using the company's model. Some of the company's economic analyses (e.g. the probabilistic sensitivity analyses) are interpreted based on the assumption that the end of life criteria are met. The ERG notes that due to the use of parametric cure models, median OS and mean OS estimates diverge significantly. The evidence available to determine whether blinatumomab satisfies NICE's end of life criteria is discussed in Chapter 6.

4. CLINICAL EFFECTIVENESS

This chapter presents a review of the clinical effectiveness evidence provided in the CS[1] for blinatumomab for treating patients with MRD+ BCP-ALL. The clinical evidence provided in the CS comprised a systematic review of randomised controlled trials (RCTs) and observational studies for adults with MRD+ BCP-ALL (Appendix D of the CS).

4.1 Critique of the methods of review

4.1.1 Searches

The company performed one clinical effectiveness search to identify all clinical and safety studies of all treatments for adult ALL patients with MRD positivity after treatment (see CS,[1] Appendix D). For the original searches undertaken in May 2017, several electronic bibliographic databases were searched including MEDLINE in Process [via PubMed], EMBASE [host not reported], the Cochrane Database of Systematic Reviews [CDSR, via Wiley Online Library], the Cochrane Central Register of Controlled Trials [CCRCT, via Wiley Online Library], the Database of Abstracts of Reviews of Effectiveness [DARE, via CRD], the NHS Economic Evaluation Database [NHS EED, via CRD] and the Health Technology Assessment database [HTA, via CRD]. Conference proceedings websites (American Society for Blood and Marrow Transplantation [ASBMT], American Society of Clinical Oncology [ASCO], American Society of Hematology [ASH], European Cancer Organisation/European Society for Medical Oncology [ECCO/ESMO], European Hematology Association [EHA], and International Society for Pharmacoeconomics and Outcomes Research [ISPOR]) were searched covering the period from June 2014 until June 2017.

According to the company's clarification response[17] (question A5), two clinical trials registers were searched on the 7th and 8th June 2017 (clinicaltrials.gov and the World Health Organization International Clinical Trials Registry Platform [WHO ICTRP]). Supplementary searches undertaken by the company also included searching unpublished Amgen studies.

For the systematic literature review searches, the company fully reported the search strategies for all the databases searched in Appendix D of the CS. The population terms comprising MeSH and free-text terms for "ALL" were combined with free-text terms for "minimal residual disease". Whilst the company have included most, if not all, of the terms for "minimal residual disease", the ERG is unable to confirm whether applying this will retrieve all ALL studies which include MRD measurement, if for example, these terms are not mentioned in the title and/or abstracts of publications.

The company applied four limits to the search strategies: (i) to exclude paediatric populations; (ii) to include only human studies; (iii) to include English language publications and (iv) to exclude certain publication types (reports, editorials, reviews, news, letters). The ERG recommends limiting the search by applying 'NOT' to exclude animal studies rather than by limiting using the 'Humans' limit function in PubMed, as the former approach is more sensitive.

The application of the English language limit suggests that the search is prone to a language bias, hence the ERG cannot confirm definitively whether any relevant non-English studies of blinatumomab have been excluded from the company's review.

No adverse event (AE) studies were identified from the searches presented in CS Appendix D. In response to a request for clarification from the ERG (see clarification response,[17] question A5), the company confirmed that a systematic search specifically for AEs for blinatumomab was not performed. The primary source of evidence on AEs was the regulatory authorities' documentation i.e. the European Public Assessment Report (EPAR). The ERG considers that the company should have undertaken a separate search for AE studies.

Aside from the issues relating to the implementation of the company's searches, the ERG considers it unlikely that any relevant studies of blinatumomab have been missed the company's searches.

4.1.2 Inclusion criteria

Appendix D of the CS describes the inclusion and exclusion criteria for acceptance into the systematic review (Table 2). One review was undertaken to identify studies of blinatumomab and its comparators (see CS, Appendix D); all studies of any interventional therapy were eligible for inclusion in the company's review.

Table 2: Eligibility criteria for the company's systematic review (reproduced from CS Appendix D Table 72)

	Inclusion criteria	Exclusion criteria
Population	Adult ALL patients with MRD positivity after treatment	• Paediatric patients • MRD- ALL patients
Intervention/comparator	Any interventional therapies	None
Outcomes	Clinical effectiveness and safety • OS • RFS • Event-free survival • MRD complete response rate • Duration of MRD response • Duration of haematologic response • Rate of transplant • Mortality following transplant • Treatment-related mortality • Serious adverse events • Grade 3 or 4 AEs (list to be determined based on the most commonly reported) • Discontinuations due to adverse events • Patient-reported outcomes	Non-clinical outcomes, such as those in pharmacodynamics or *in vitro* studies
Study design	• RCTs of at least 10 patients per arm • Single-arm clinical trials of at least 10 patients • Prospective and retrospective observational studies of at least 10 patients	• Case studies and studies evaluating fewer than 10 patients • Letters, narrative reviews, expert opinions, etc.

As stated in the decision problem (see Table 1), the comparator of "monitor for relapse" that was specified in the final NICE scope[22] was not considered in the CS. The clinical advisors to the ERG noted that some older and less fit patients may not be able to receive HSCT or to tolerate chemotherapy; however, they may be able to tolerate blinatumomab. The clinical advisors noted that this population would be small. The ERG considers that the exclusion of this comparator is not appropriate, although no clinical evidence is reported for the use of blinatumomab in this subgroup. It is unclear whether alternative sources (for example, unpublished registry data) may have provided evidence for this comparator.

The included population relates to adult ALL patients with MRD positivity after treatment. Different technologies could be used to define MRD positivity (multicolour flow cytometry to detect abnormal immunophenotypes; real-time quantitative polymerase chain reaction [RT-qPCR] assays to detect clonal rearrangements in Ig heavy chain genes, and/or T-cell receptor [TCR] genes; RT-qPCR assays

to detect fusion genes) (CS Section B.1.3). This was not an inclusion criterion, but was recorded for each study. The ERG considers this to be appropriate.

Included outcomes were AEs, patient-reported outcomes (PROs), and the following clinical effectiveness outcomes: OS; RFS; event-free survival (EFS); MRD complete response rate; duration of haematologic response; rate of transplant, and mortality following transplant. These outcomes are consistent with the final NICE scope,[22] with the addition of duration of haematologic response which was not listed in the final NICE scope.

Study designs eligible for inclusion in the company's review included RCTs, single-arm studies and prospective and retrospective observational studies. The ERG considers this to be appropriate. Studies were only included if they had at least 10 patients (or 10 per arm for RCTs). The ERG considers this criterion to be arbitrary and notes that its application could lead to the exclusion of small but relevant studies. However, following a request for clarification from the ERG (see clarification response,[17] question A6) the company confirmed that no relevant studies were excluded for this reason. Included publications were limited by English language, but not location of study.

Study selection was conducted by two independent reviewers, with disagreements resolved by a third reviewer, in accordance with good practice for systematic reviews.

4.1.3 Critique of data extraction

The company's clarification response[17] (question A7) states that data extraction was conducted independently by two reviewers, with disputes resolved by a third reviewer. The ERG considers this to reflect good practice in systematic reviews.

Based on the information provided in Section B.2 of the CS, it was apparent that relevant data were extracted on study methodology and patient characteristics. CS Appendix D explicitly states that data were extracted on definition of MRD and subgroups according to CR status after first-line or salvage treatment (see CS, Table 73).

Data extracted for the three studies and included in the CS (see Section 4.2) were checked by the ERG against clinical study reports (CSRs) and were found to be accurate.

4.1.4 Quality assessment of included studies

The company's quality assessment was conducted independently by two reviewers, with disputes resolved by a third reviewer (see clarification response,[17] question A7), as is good practice in systematic reviews.[27]

Quality assessment of the two included blinatumomab studies (BLAST and MT103-202), and the retrospective study of standard care chemotherapy (Study 20120148), was presented in CS Appendix D (see CS,[1] Table 74). The quality assessment tool used in the CS was the Risk Of Bias In Non-randomised Studies of Interventions (ROBINS-I) checklist.[28] This checklist is designed to assess risk of bias within a given study, but does not address external validity, and therefore does not address issues relating to generalisability or the limitations of particular study designs.[28] The three included studies were all single-arm, open-label studies; the CS acknowledges that non-randomised study designs are subject to limitations. Based on the ROBINS-I checklist, the company deemed the overall risk of bias of all three studies to be low.

The ROBINS-I checklist is designed for non-randomised studies of interventions *"that compare the health effects of two or more interventions"* (detailed guidance is available from https://sites.google.com/site/riskofbiastool//welcome/home). Whilst there is no universally accepted validated tool for critically appraising single-arm studies, several checklists have been developed and applied to case series.[29] The ERG assessed the quality of the single-arm studies based on the criteria for case series suggested by the Centre for Reviews and Dissemination (CRD, see Table 3).[30,29]

Table 3: Quality assessment of the three included studies

CRD criteria	BLAST[31]	MT103-202[32]	Study 20120148[33]
Is the study based on a representative sample selected from a relevant population?	Yes	Yes	Yes
Are the criteria for inclusion explicit?	Yes	Yes	Yes
Did all individuals enter the survey at a similar point in their disease progression?	Yes	Yes	Yes
Was follow-up long enough for important events to occur?	Yes	Yes	Yes
Were outcomes assessed using objective criteria or was blinding used?	Outcome assessors were not blinded OS – objective criteria MRD response– objective criteria RFS – objective criteria HRQoL – at risk of bias	Outcome assessors were not blinded MRD – objective criteria RFS – objective criteria	Outcome assessors were not blinded OS – objective criteria RFS – objective criteria

The studies were well conducted according to CRD criteria.[30] Prognostic factors such as disease stage and age were reported in the CS for all three included studies. Statistical analyses including subgroup analyses were pre-specified.[31-33] In the BLAST study there was a low risk of attrition bias, all patients included at baseline and treated with at least one cycle of blinatumomab were included in the OS analysis, and the majority of these patients were included in the RFS (95%) and MRD (97%) analyses. MT103-202 included all 20 patients treated with at least one cycle of blinatumomab in MRD and RFS analyses.

Single-arm studies are low on the hierarchy of study quality as they are associated with potential biases.[34] The absence of blinding leads to a risk of performance bias.[27, 30] The lack of randomisation leads to a risk of selection bias.[27, 30]

Eligibility criteria for all three included studies were adequately described in the CS. However, it was not clear from the CS how patients were identified for recruitment into the blinatumomab studies and whether patients were recruited consecutively.[29, 35] The company's clarification response[17] (question A16) provides reasons for not enrolling 95 screened patients, most of which were due to patients not meeting eligibility criteria for MRD level ($< 1 \times 10^{-3}$) or having an overt relapse.

Single-arm studies also have a risk of detection bias due to the absence of blinding. One means by which the risk of bias can be reduced in open-label studies is to introduce blinded outcome assessors. However, in the included studies, the lack of blinding is unlikely to impact on OS, MRD or RFS. The HRQoL outcome is necessarily prone to bias as it is patient-reported and therefore assessor-blinding is not possible in an open-label study.[36]

Study 20120148 comprised a retrospective analysis of existing clinical databases. Retrospective studies are more likely to be susceptible to bias than prospective studies, particularly selection bias.[30, 37] However, Study 20120148 was used in the CS to select a population comparable to that of the BLAST study, rather than to provide a population representative of all MRD+ BCP-ALL patients.

4.1.5 Evidence synthesis

Due to the lack of RCT data to inform the comparative effectiveness of blinatumomab versus standard care chemotherapy, the company synthesised data from BLAST and Study 20120148 using indirect comparison methods. Further details of this analysis are provided in Section 4.4.

4.2 Included blinatumomab studies

The CS[1] included three studies identified by the systematic searches. Two studies were of blinatumomab (MT103-202 and BLAST). The third study was a historical comparator study (Study 20120148, described in Section 4.3). All three studies were sponsored by Amgen and information was provided in CSRs[31,32,33] and the BLAST protocol.[38] At the time of writing, BLAST was published in two abstracts.[39, 40]

No relevant RCTs were identified by the company or by the ERG. The ERG does not believe that any relevant studies of blinatumomab retrieved from the searches were excluded from the CS.

An ERG search of the U.S. National Institutes of Health (NIH) clinical trials registry[41] identified two potentially relevant ongoing studies, however, at the time of writing, the completion dates for these studies were more than 12 months in the future. Study NCT02458014 (Blinatumomab in Patients with B-cell Lineage Acute Lymphocytic Leukaemia with Positive Minimal Residual Disease) has an estimated primary completion date of September 2020. Study NCT02767934 (Pembrolizumab in Treating Minimal Residual Disease in Patients with Acute Lymphoblastic Leukaemia), which includes B-cell as well as T-cell ALL, has an estimated primary completion date of January 2019.

4.2.1 Study characteristics of blinatumomab studies

The two included studies of blinatumomab (MT103-202[32] and BLAST[31]) were both single-arm studies. Study characteristics are shown in Table 4. Within the limitations of the study design, the studies were well conducted, however, as noted in Section 4.1.4, there are biases associated with single-arm studies.

Table 4: Characteristics of included blinatumomab studies

Study	Reference(s)	Study design	Population	Number enrolled	Intervention	Primary outcome	Dates of enrolment	Follow-up
MT103-202 NCT00560794	Amgen CSR 2013[32]	Phase II, single-arm, open-label, multicentre	Adult MRD+ BCP-ALL patients in haematological CR after front-line therapy	[redacted] 20 received at least one cycle and included in efficacy analysis (of 21 who received at least one infusion and were included in safety analysis)	Blinatumomab 15µg/m^2/day continuous infusion. Up to 10 cycles	Incidence of MRD negativity/response** within 4 cycles of treatment with blinatumomab	2008 – 2009	Primary efficacy 24 weeks (4 cycles) Safety up to 4 weeks after last treatment
BLAST MT103-203 NCT01207388	Amgen CSR 2016[31] Amgen protocol 2010[38] Goekbuget 2014[39]	Phase II, single-arm, open-label, international, multicentre	Adult MRD+ BCP-ALL patients in haematological CR after front-line therapy	[redacted], 116 received at least one infusion	Blinatumomab 15µg/m^2/day continuous infusion* Up to 4 cycles	Proportion of patients who achieve complete MRD response defined by absence of MRD after one cycle of treatment	2010 – 2014	Safety 30-days Efficacy 9, 12, 18, and 24 months Survival 30, 36, 42, 48, 54, 60 months

Information from CS Section B.2.2 and B.2.3 and Appendix D, CSRs,[31, 32] and U.S. National Institutes of Health clinical trials registry[41]
** one cycle = continuous infusion for four weeks followed by two-week infusion-free interval*
*** MRD negativity/response defined as bcr/abl and/or t[4;11] below detection limit and/or individual rearrangements of immunoglobulin or TCR genes below 10^{-4} [31] bcr/abl = "breakpoint cluster region/gene on human chromosome #9" [31]*

Eligibility criteria

BLAST inclusion criteria

Patients were eligible for inclusion in the study only if all the following criteria applied:

- Patients with BCP-ALL in complete haematological remission defined as less than 5% blasts in bone marrow after at least three intense chemotherapy blocks (e.g., GMALL induction I–II/consolidation I, induction/intensification/consolidation or three blocks of Hyper CVAD)
- Presence of MRD at a level of $\geq 10^{-3}$ (molecular failure or molecular relapse) in an assay with a sensitivity and a lower level of quantification of 10^{-4} documented after an interval of at least 2 weeks from last systemic chemotherapy
- For evaluation of MRD, patients must have had at least one molecular marker based on individual rearrangements of immunoglobulin or TCR-genes or a flow cytometric marker profile evaluated by a national or local reference lab approved by the sponsor
- Bone marrow specimen from primary diagnosis (enough DNA [30pg] or a respective amount of cell material) for clone-specific MRD assessment must have been received by central MRD lab and lab must have confirm that the sample is available
- Bone marrow function as defined below:
    - ANC (Neutrophils) ≥1,000/µL
    - Platelets ≥50,000/µL (transfusion permitted)
    - Haemoglobin (HB) level ≥9g/dI (transfusion permitted)
- Renal and hepatic function as defined below:
    - AST (GOT), ALT (GPT), and AP <2 x ULN
    - Total bilirubin <1.5 x ULN
    - Creatinine clearance ≥50 mL/min (calculated e.g. per Cockroft & Gault)
- Negative HIV test, negative hepatitis B (HbsAg) and hepatitis C virus (anti-HCV) test
- Negative pregnancy test in women of childbearing potential
- Eastern Cooperative Oncology Group (ECOG) Performance Status 0 or 1
- Age ≥18 years
- Ability to understand and willingness to sign a written informed consent
- Signed and dated written informed consent.

BLAST exclusion criteria

Patients were excluded from participation in the study if any of the following criteria applied:

- Presence of circulating blasts or current extra-medullary involvement by ALL
- History of relevant central nervous system (CNS) pathology or current relevant CNS pathology (e.g. seizure, paresis, aphasia, cerebrovascular ischemia/haemorrhage, severe brain injuries,

dementia, Parkinson's disease, cerebellar disease, organic brain syndrome, psychosis, coordination or movement disorder)

- Current infiltration of cerebrospinal fluid (CSF) by ALL
- History of or active relevant autoimmune disease
- Prior allogeneic HSCT
- Eligibility for treatment with TKIs (i.e., Ph+ patients with no documented treatment failure of or intolerance/contraindication to at least 2 TKIs)
- Systemic cancer chemotherapy within 2 weeks prior to study treatment (except for intrathecal prophylaxis)
- Radiotherapy within 4 weeks prior to study treatment
- Autologous HSCT within six weeks prior to study treatment
- Therapy with monoclonal antibodies (rituximab, alemtuzumab) within 4 weeks prior to study treatment
- Treatment with any investigational product within four weeks prior to study treatment
- Previous treatment with blinatumomab
- Known hypersensitivity to immunoglobulins or to any other component of the study drug formulation
- History of malignancy other than ALL within five years prior to treatment start with blinatumomab, except for basal cell or squamous cell carcinoma of the skin, or carcinoma "in situ" of the cervix
- Active infection, any other concurrent disease or medical condition that are deemed to interfere with the conduct of the study as judged by the investigator
- Nursing women or women of childbearing potential not willing to use an effective form of contraception during participation in the study and at least 3 months thereafter or male patients not willing to ensure effective contraception during participation in the study and at least three months thereafter.

Study MT103-202 eligibility criteria (from CS[1] Section B.2.3 and the NIH clinical trials registry[41]) were as follows:

Study MT103-202 inclusion criteria

- Adults (≥18 years of age) with BCP-ALL
- MRD positivity at a level of at least $1x10^{-4}$ at any point after the first consolidation chemotherapy block of front-line therapy
- ECOG Performance Status < 2.

Study MT103-202 exclusion criteria

- Current extramedullary involvement
- History of (or current) clinically relevant CNS pathology or autoimmune disease
- Prior autologous HSCT (within 6 weeks) or allogeneic HSCT (at any time)
- Chemotherapy or radiotherapy (within 4 weeks)
- Therapy with monoclonal antibodies (within 6 weeks)
- Known hypersensitivity to immunoglobulins or to any other component of the study drug formulation
- History of malignancy other than ALL within five years prior to treatment start with blinatumomab, except for basal cell or squamous cell carcinoma of the skin, or carcinoma "in situ" of the cervix
- Active infection, any other concurrent disease or medical condition that are deemed to interfere with the conduct of the study as judged by the investigator
- Nursing women or women of childbearing potential not willing to use an effective form of contraception during participation in the study and at least 3 months thereafter or male patients not willing to ensure effective contraception during participation in the study and at least three months thereafter.

Table 5 presents the baseline characteristics of BLAST and MT103-202. The studies had similar baseline ages (BLAST median age=45 years, MT103-202 mean age=50 years). The majority of participants were Ph-, with n=[redacted] Ph+ participants in each study. In the BLAST study, the majority of patients (65%) were in first CR. The demographics of the study were considered by clinical advice to be similar to those of the UK population with BCP-ALL who have MRD activity while in remission. In the BLAST study, the majority of patients (84%) had a baseline MRD level of between 10^{-3} and 10^{-1}, where patients are classed as MRD+ with disease measurable to 10^{-4}. These MRD levels may not necessarily reflect those of the UK population, but reflect the eligibility criteria for the blinatumomab studies.

Table 5: Baseline characteristics of participants in BLAST and MT103-202

Commented [LM1]: CHECK ACIC

Baseline characteristic	**BLAST MT103-203 (n=116)**		**MT103-202 (Pilot) (n=20)**	
Male sex, n (%)	68 (59)		8 (40.0)	
Age, years	Median (range)	45.0 (18–76)	Mean (SD)	49.8 (18.3)
Age, n (%)	≥18 to <35 years	36 (31.0)	20-30 years	3 (15.0)
	≥35 to <55 years	41 (35.3)	31-40 years	5 (25.0)
	≥55 to <65 years	24 (20.7)	41-50 years	2 (10.0)
	≥65 years	15 (12.9)	51-60 years	1 (5.0)
			61-70 years	7 (35.0)
			> 70 years	2 (10.0)
Median time from prior treatment (range), months	[redacted]		NR	
Relapse history, n (%)	First CR	75 (65)	NR	
	Second CR	39 (34)		
	Third CR	2 (2)		
Baseline MRD levels at cetral laboratory, n (%)	$\geq 10^{-1} < 1$	9 (7.8)	NR	
	$\geq 10^{-2} < 10^{-1}$	45 (38.8)		
	$\geq 10^{-3} < 10^{-2}$	52 (44.8)		
	$< 10^{-3}$	3 (2.6)		
	Below LLQ	5 (4.3)		
	Unknown	2 (1.7)		
Philadelphia chromosome disease status, n (%)	Positive	5 (4.3)	Positive	[redacted] (CS Clarification response[17] A2)
	Negative	111 (95.7)	Negative	[redacted]
Ethnicity, n (%)	White: 102 (87.9) Asian: 1 (0.9) Mixed: 1 (0.9) Unknown: 12 (1.3)		Caucasian 20 (100.0)	
Genetic alterations, n (%)	Confirmed t(4;11) Translocation / MLL-AF4+	5 (6.8)	Confirmed t(4;11) Translocation / MLL-AF4+	2 (10.0)
			bcr/abl above detection limit (all)	5 (25.0)
WBC at diagnosis (>30,000/mm^3), n (%)	18 (15.5)		NR	

Information from CS Section B.2.3, Goekbuget 2014[39] and U.S. National Institutes of Health clinical trials registry[41]
Values in parentheses represent percentages
CR - complete response; LLQ - lower limit of quantification; WBC - white blood cell; NR - not reported

Prior ALL treatment received by patients in the BLAST study was provided in the company's clarification response[17] (question A12), and is shown in Table 6. Prior front-line treatment had been received by [redacted] of the patients, [redacted] had received treatment for first relapse, and only [redacted] had received treatment for second relapse. For prior anti-tumour drug treatment, [redacted]

[redacted]

[redacted] Clinical advice received by the ERG suggested that these treatments would lead to most patients in the study having a similar level of disease to those patients seen in current practice who are eligible for treatment using blinatumomab in England. Although PETHEMA is quite different from practice in England, only a small percentage of patients ([redacted]) received this regimen.

Table 6: BLAST study Prior ALL treatment (reproduced from company's clarification response question A12)

Characteristic Category	Full analysis set (n=116)
Maximum line of therapy	
Front-line treatment	[redacted]
First relapse treatment	[redacted]
Second relapse treatment	[redacted]
Front-line treatment	[redacted]
Pre-phase	[redacted]
GMALL	[redacted]
combination of regimen /other	[redacted]
GMALL elderly	[redacted]
GRAALL	[redacted]
UKALL	[redacted]
GIMEMA	[redacted]
PETHEMA	[redacted]
FLAG-Ida	[redacted]
NILG	[redacted]
TKI	[redacted]
FRAALLE	[redacted]
Hyper-CVAD	[redacted]
iBFM	[redacted]
AIEOP	[redacted]
HOVON	[redacted]
ALL-2009	[redacted]
ALL-2009 elderly	[redacted]
EWALL elderly	[redacted]
GRAAPH	[redacted]
LALA94	[redacted]
Romanian Group for ALL	[redacted]

Concomitant medications allowed are shown in Table 7. Clinical advice received by the ERG suggested that these are similar to current practice in England.

Table 7: Concomitant medications allowed in the blinatumomab studies (data extracted from CS Table 12)

	BLAST	**MT103-202**
Permitted medications	Prior to the start of cycle 1: CSF (cerebrospinal fluid) prophylaxis A corticosteroid Prior to the start of subsequent cycles: A corticosteroid During the treatment period: Dexamethasone in the case of neurologic events Following treatment cycles 2 and 4 immediately after bone marrow aspiration: CSF prophylaxis After completion of study treatment for patients who did not undergo HSCT: CSF prophylaxis Patients at high risk for CMV infection: Intensive CMV-PCR follow-up or prophylactic CMV treatment	Premedication for each treatment cycle included a corticosteroid to suppress cytokine release (100mg methylprednisolone IV at 1 hour prior to start of blinatumomab infusion or prior to restart if infusion interruption > 12 hours) and thrombosis prophylaxis by low molecular weight heparin (subcutaneous) during the first 7 days of each treatment cycle CNS prophylaxis was administered with the following intrathecal triple combination regimen at absolute doses: dexamethasone 4mg, methotrexate 15mg, cytosine-arabinoside 40mg. If the patient had MRD response after cycle 1 of treatment, the triple combination regiment was administered immediately after the first bone marrow aspiration study on day 28 of cycle 2 In non-responders, after cycle 1 demonstrated detectable MRD, the triple combination regimen was administered after cycle 3 of treatment immediately after bone marrow aspiration on cycle day 28 of cycle 3. CNS prophylaxis continued every 3 months Small molecule TKIs registered for the treatment of ALL disease were permitted as concomitant treatment of patients with bcr/abl positive MRD if the patients developed MRD relapse on TKIs or whose MRD persisted on TKIs for more than 8 weeks For symptomatic treatment of fever, metamizole was administered
Disallowed medications	Any anti-tumour therapy Any other investigational agent Chronic systemic high-dose corticosteroid therapy Any other immunosuppressive therapies Non-steroidal anti-inflammatory drugs Paracetamol/acetaminophen was allowed TKIs	Any anti-tumour therapy other than blinatumomab as indicated in the protocol Any other investigational agent Chronic systemic high-dose corticosteroid therapy Other immunosuppressive therapies Stem-cell transplantation Any use of NSAIDs (nonsteroidal anti-inflammatory drugs) except for paracetamol

4.2.2 Clinical effectiveness in the blinatumomab studies

Overall Survival

Study MT103-202 did not include OS as an outcome measure. OS data from BLAST are shown in Table 8 and Figure 2. OS was defined as the time from first blinatumomab treatment until death due to any cause, with patients who did not die being censored at their last contact date (see CS,[1] Section B.2.6.1).

Results from Cox proportional hazards models including treatment as a covariate in the model are presented by the company. Results from the primary analysis, described by the company as "without censoring at HSCT" are used for the health economic model. Results are also presented including a time-dependent covariate for HSCT, described by the company as "with censoring at HSCT"; the company states that this analysis was conducted to account for differences between transplant rates in BLAST and the historical control, hence it better isolates the blinatumomab treatment effect not affected by use of transplant.[1] The primary analysis (without censoring at HSCT) is considered by the ERG to be most appropriate as, according to the CS, blinatumomab is not intended to displace HSCT in the treatment pathway.

After a median follow-up of 18 months, median OS was [redacted]

Figure 2: OS in BLAST (reproduced from CS Figure 13)

[redacted]

Subgroup analysis of OS outcomes was conducted according to the following factors: age; gender; Philadelphia status; patients by t(4;11) translocation and/or MLLAF4+ ALL haematological remission; risk stratification; relapse history; MRD level at baseline; white blood cells (WBC) at first diagnosis; chemo-resistance after the first week of chemotherapy; need of salvage therapy for CR; previous anti-

tumour radiotherapies; incidence of neurologic events during cycle 1; time from diagnosis to start of blinatumomab; time from last treatment to start of blinatumomab, and clinical trial material from manufacturing process 4/5. The only subgroup which was found to differ significantly for OS was Ph status, with Ph- patients experiencing a significantly [redacted] median OS than Ph+ patients ([redacted]) (see CS,[1] Section B2.7). This was based on only 5 Ph+ patients, all of whom were in CR2/3 rather than CR1, hence the ERG considers that the interpretation of this subgroup finding should be treated with caution.

Table 8: Summary of OS outcomes in BLAST

	BLAST (n=116)	
Outcome	**OS not censored at HCST**	**OS censored at HCST**
Events, n (%)	[redacted]	[redacted]
Censors, n (%)	[redacted]	[redacted]
OS % (18 months)	[redacted]	[redacted]
95% CI	[redacted]	[redacted]
Median (months)	[redacted]	[redacted]

Information from CS Section B.2.6.1 (Table 21 of the CS) and U.S. National Institutes of Health clinical trials registry[41]
n.e. = not estimable

Relapse-free survival

Haematological RFS in BLAST was measured from the first dose of blinatumomab until the first assessment of documented relapse (either haematological (>5% leukaemia cells in bone marrow as measured by cytological, microscopic assessment, presence of circulating leukaemia blasts) or extramedullary leukaemia), secondary leukaemia, or death due to any cause.[41] In the MT103-202 study, time to haematological relapse was defined as the time between the start of first infusion of blinatumomab and the first result of haematological relapse, defined as >5% leukaemia cells in bone marrow.[41]

RFS data were provided by 110 patients in the BLAST study who were in haematological CR at baseline, excluding Ph+ participants (see Table 9) as Ph+ patients were excluded from the pre-specified secondary analyses.[41] At 18-months follow-up, the uncensored median time to haematological relapse was [redacted]. In the MT103-202 study, RFS was [redacted] at five years.

Figure 3: RFS in BLAST (reproduced from CS Figure 10)

Subgroup analysis of RFS outcomes was conducted according to the following factors: age; gender; Philadelphia status; patients by t(4;11) translocation and/or MLLAF4+ ALL haematological remission; risk stratification; relapse history; MRD level at baseline; WBC at first diagnosis; chemo-resistance after the first week of chemotherapy; need of salvage therapy for CR; previous anti-tumour radiotherapies; incidence of neurologic events during cycle 1; time from diagnosis to start of blinatumomab; time from last treatment to start of blinatumomab, and clinical trial material from manufacturing process 4 or 5. The only subgroup found to differ significantly was relapse history.

(see CS,[1] Section B2.7).

Table 9: Summary of RFS outcomes in BLAST and MT103-202

Study	**BLAST n=110**		**MT103-202 (Pilot) n=20**
Outcome	**RFS not censored at HCST**	**RFS censored at HCST**	
Events, n (%)			NR
Censors, n (%)			NR
RFS %			
95% CI			
Median RFS			

(months)			

Information from CS Section B.2.6.1 (Table 20 and Figure 10 of the CS) and CS Section B.2.6.2

Minimal residual disease response

Within the BLAST study, complete MRD response was defined as no polymerase chain reaction (PCR) amplification of individual rearrangements of immunoglobulin (Ig)- or TCR -genes detected (the minimum required sensitivity of 1 x 10^{-4}) after completion of the first cycle (see CS Section B.2.3 and US NIH clinical trials registry[41]). [redacted][38]

For patients in the MT103-202 study, the primary endpoint was MRD response rate within four cycles of blinatumomab. For Ph+ or translocation (t) (4;11) patients, response was achieved when Ph or t(4;11) was below detection limit and individual rearrangements of immunoglobulin or TCR genes were below 1 x 10^{-4}. For Ph- and t(4;11) negative, response was achieved when individual rearrangements of immunoglobulin or TCR genes were below 1 x 10^{-4} (see CS,[1] Section B.2.3 and US NIH clinical trials registry[41]). MRD response outcomes are presented in Table 10. All 80% of patients achieving MRD response did so within one cycle.

In the BLAST study, three patients were excluded from the MRD response analysis due to missing data (n=1) or assays with a sensitivity of 5×10^{-4} (n=2).[39] Data on MRD response from 113 patients in BLAST are shown in Table 10. A total of ninety patients ([redacted]) achieved MRD response after one or more cycles of blinatumomab treatment, with 88 of these patients responding within one cycle.[39] There was a higher rate of response for patients in first CR 82% (95% CI 72% to 90%), than in second CR 71% (95% CI 54% to 85%) or third CR 50% (95% CI 1% to 99%); however, only two patients were in third CR (see Table 5), hence results on this subgroup should be treated with caution.[39] For other subgroups, there was no significant difference in MRD response (age, sex, line of treatment, and MRD levels).[39]

Table 10: MRD response in the blinatumomab studies

Outcome	BLAST [redacted]	MT103-202 (Pilot) (n=20)
Complete MRD response after 1 cycle n (%,95% CI)	[redacted]	[redacted]
Complete MRD response after ≥1 cycle n (%, 95% CI)	[redacted]	[redacted]
Median time to MRD response, days	[redacted]	NR
Duration of median response, months (uncensored)	[redacted]	[redacted]).
Duration of median response, months (censored at HCST or post-blinatumomab chemotherapy)	[redacted]	NR

Information from CS Section B.2.6.1 (Tables 18 and 19 of the CS) and CS Section B.2.6.2 and Goekbuget 2014.[39]
**participants who were in haematological complete remission at treatment start, excluding Ph+, who had an MRD complete response at cycle 1*

Rate of stem cell transplant

For Study MT103-202 [redacted] patients underwent HCST (see CS,[1] Section B.2.6.2). In BLAST, [redacted] patients underwent HCST, of whom [redacted] were in complete haematological CR at the time of HSCT. Within the group of 74 Ph- patients who underwent HSCT prior to relapse, the 100-day mortality probability was 7%.[1]

Health-related quality of life

BLAST measured HRQoL using the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire C30 (EORTC QLQ-C30),[42] a validated, cancer-specific patient reported outcome questionnaire, and the Euroqol 5-Dimensions questionnaire (EQ-5D, available from: https://euroqol.org/), a standardised measure of generic health status. HRQoL results from BLAST are shown in Table 11 and Table 12, respectively. [redacted] HRQoL as measured by EQ-5D did not change significantly from baseline to the end of the core study.

Table 11: Change from baseline in EORTC-QLQ-C30 scales in BLAST (n=116) (reproduced from CS Table 24)

EORTC-QLQ-C30 Scale	Baseline, mean (SE) (Max=100)	Greatest change from baseline in cycles 1 to 4, mean (SE)/cycle	Change from baseline at end of core study, mean (SE)
Global health status	[redacted]	[redacted]	3.9 (2.4)
Physical function	[redacted]	[redacted]	2.2 (1.9)
Role functioning	[redacted]	[redacted]	1.4 (3.5)
Emotional functioning	[redacted]	[redacted]	5.3 (2.7)
Cognitive functioning	[redacted]	[redacted]	-2.3 (2.5)
Social functioning	[redacted]	[redacted]	14.9 (3.8)
Fatigue	[redacted]	[redacted]	-5.4 (2.4)
Nausea and vomiting	[redacted]	[redacted]	-2.3 (2.0)
Pain	[redacted]	[redacted]	-1.4 (2.7)
Dyspnoea	[redacted]	[redacted]	-0.9 (2.9)
Insomnia	[redacted]	[redacted]	3.7 (3.5)
Appetite loss	[redacted]	[redacted]	-9.1 (3.4)
Constipation	[redacted]	[redacted]	0 (2.2)
Diarrhoea	[redacted]	[redacted]	0.0 (2.3)
Financial difficulties	[redacted]	[redacted]	-0.9 (2.9)

EORTC-QLQ-C30 - European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-C30; SE - standard error

Table 12: Change from baseline in EQ-5D domains in BLAST (n=116) (reproduced from CS Table 25)

EQ-5D scale	Baseline, mean (SE)	Greatest change from baseline in cycles 1 to 4, mean (SE)/cycle	Change from baseline at end of core study, mean (SE)
Mobility	[redacted]	[redacted]	0 (0.1)
Self-care	[redacted]	[redacted]	0 (0.0)
Usual activity	[redacted]	[redacted]	-0.1 (0.1)
Pain/discomfort	[redacted]	[redacted]	-0.1 (0.1)
Anxiety/depression	[redacted]	[redacted]	-0.1 (0.1)

EQ-5D - EuroQol 5-dimensions; SE - standard error

Adverse events

Differences in treatment regimen for BLAST and MT103-202

The CS[1] (Section B.2.10.1) reports both treatment-emergent and treatment-related (considered to be related to blinatumomab) AEs for MRD+ BCP- ALL patients. Pooled data from the BLAST study[31] (n=116) and MT103-202[32] (n=21) are reported in the CS.[1] A meta-analysis of data from the two studies was not conducted. The CSR for BLAST[31] reports a median treatment duration of 55 days, whilst the CSR for MT103-202[32] reports a median treatment duration of 87.3 days. The CSRs report that the

timeframe for recording of AEs was from the first dose of blinatumomab to 30 days after the last dose for BLAST, and from the first dose of blinatumomab to 4 weeks after the last dose for MT103-202. For BLAST, the CSR reports a dosing regimen of blinatumomab continuous IV infusion at 15µg/m^2/day at a constant flow rate over 28 days, followed by an infusion-free interval of 14 days, for up to 4 cycles.[31] For MT103-202, the CSR reports a dosing regimen of blinatumomab continuous IV infusion at 15µg/m^2/day at a constant flow rate over 28 days, followed by an infusion-free interval of 14 days, for up to 7 cycles in patients who showed neither MRD progression nor response.[32] Patients who had achieved MRD response were administered 3 additional cycles of treatment, up to a maximum of 10 cycles. For MT103-202, a dose increase to 30µg/m^2/day was permitted where there was evidence of insufficient response.[32]

Numbers of adverse events (BLAST and MT103-202)

A summary of the pooled treatment-emergent and treatment-related AEs for MRD+ BCP-ALL patients from BLAST and MT103-202 is presented in Table 13. Amongst MRD+ BCP-ALL patients, [redacted] participants experienced at least one treatment-emergent AE. [redacted] of participants experienced an AE classed as serious; [redacted] of patients had Grade ≥3 AEs; and [redacted] of patients had Grade ≥4 AEs. [redacted] participants experienced treatment-emergent AEs that led to the discontinuation of treatment; [redacted] of participants had a serious adverse event (SAE); [redacted] of patients had Grade ≥3 AEs; and [redacted] of patients had Grade ≥4 AEs. [redacted] suffered AEs that were considered to be treatment-related. [redacted] of participants experienced a treatment-related AE that was classed as serious; [redacted] that were classed as Grade ≥3; and [redacted] that were classed as Grade ≥4. [redacted] participants [redacted] experienced treatment-related AEs that led to the permanent discontinuation of treatment. [redacted] of participants had a serious event; [redacted] Grade ≥3; [redacted] Grade ≥4.

The most common treatment-emergent events of interests (EOIs) were [redacted] of participants); [redacted], and [redacted]. Other EOIs with a frequency of ≥5% were [redacted]; [redacted]; [redacted]; [redacted]; and [redacted].

There was [redacted] which was considered related to blinatumomab, and [redacted] which was not considered to be treatment-related.

A summary of disaggregated data reporting the frequency of SAEs for BLAST and MT103-202 is presented in Table 14. These data were taken from the US NIH clinical trials registry and cross-

referenced against both CSRs (BLAST CSR Table 14-6.25; MT103-202 Table 12-4). [redacted] of patients in BLAST and [redacted] of patients in MT103-202 experienced an SAE. The most common of these were blood and lymphatic system disorders, infections and infestations, injury, poisoning or procedural complications, and nervous system disorders. In BLAST,[31] [redacted] patients experienced SAEs classified as general disorders. The most common of these was pyrexia. No reports of SAEs classified as general disorders are reported in the CSR for MT103-202.[32] In addition, in BLAST, [redacted] patients experienced SAEs relating to investigations, whilst none were reported in MT103-202.

The CS[1] draws comparisons between the pooled data from BLAST and MT103-202 and the known safety profile of blinatumomab in adult patients with relapsed or refractory Ph-BCP-ALL. This profile comprised pooled data from MT103-206, MT103-211, and TOWER. Table 15 presents data for AEs for these two sets of pooled data, as reported in the CS. Safety profiles are consistent between these populations, with the following exceptions: (i) treatment-emergent Grade ≥3 AEs were lower in MRD+BCP-ALL patients ([redacted]); (ii) there was a higher rate of treatment-related AEs for the MRD+BCP-ALL population ([redacted]), and (iii) there was a difference for treatment-related SAEs, which were higher in MRD+BCP-ALL patients ([redacted]), although the CS reports that this is likely due to a high rate of Grade ≥2 AEs. Grades ≥3 and 4 AEs were comparable between the two populations ([redacted]). With respect to treatment-emergent EOIs, whilst there was a comparable rate of any-grade EOIs between populations, a lower rate of EOIs Grade ≥3 and Grade ≥4 was reported in the MRD+BCP-ALL population ([redacted], respectively). Both populations experienced similar rates of neurological AEs ([redacted]). A lower rate of cytokine release syndrome was reported in MRD+BCP-ALL patients compared with Ph-BCP-ALL patients ([redacted]). The CS suggests this may be a result of a lower disease burden in the MRD+BCP-ALL population in haematological CR. Rates of treatment interruptions were consistent between both populations.

The CS[1] reports that the safety profile of blinatumomab in adult MRD+ BCP-ALL patients reflects its known safety profile in a Ph-BCP-ALL population, with no new risks suggested. The blinatumomab EPAR[43] including AE data for the MRD+BCP-ALL population was unavailable at the time of writing (EMA accessed 30th January 2018), but is expected to be published early 2018 (see CS,[1] Appendix C).

Table 13: Incidence of treatment-emergent and treatment-related AEs from pooled data from the BLAST study and MT103-202 for MRD+ BCP-ALL (adapted from CS Table 30)

Event	Treatment-emergent AEs	Treatment-related AEs
All AEs, n (%)		
Serious		
Grade ≥3		
Grade ≥4		
Fatal*		
Leading to permanent discontinuation of blinatumomab		
Serious		
Grade ≥3		
Grade ≥4		
Fatal*		
Leading to interruption of blinatumomab		
Serious		
Grade ≥3		
Grade ≥4		
Fatal*		

** Fatal events that occurred within 30 days of last blinatumomab treatment*

Table 14: SAEs in MRD+ BCP-ALL patients for BLAST and MT103-202

SAE	BLAST MT103-203 (n=116)		MT103-202 (Pilot) (n=21)	
SAEs	73		10	
Blood and lymphatic system disorders		Anaemia (1); bone marrow failure (1); febrile neutropenia (2); leukopenia (1); neutropenia (5); thrombocytopenia (1)		Leukopenia (1); lymphopenia (6); thrombocytopenia (1)
Cardiac disorders		Sinus bradycardia (1); sinus tachycardia (1)		NR
Gastrointestinal disorders		Abdominal pain (1); diarrhoea (1); gastrointestinal haemorrhage (1)		NR
General disorders		Device issue (1); device malfunction (2); fatigue (1); gait disturbance (1); infusion site extravasation (1); product contamination microbial (1); puncture site pain (1); pyrexia (17); thrombosis in device (1)		NR
Hepatobiliary disorders		Hepatotoxicity (1)		NR
Immune system disorders		Cytokine release syndrome (2); hypersensitivity (2)		NR

SAE	BLAST MT103-203 (n=116)		MT103-202 (Pilot) (n=21)	
Infections and infestations	[redacted]	Acinetobacter bacteraemia (1); atypical pneumonia (1); bacterial infection (1); bronchopneumonia (1); bronchopulmonary aspergillosis (1); cystitis klebsiella (1); device related infection (3); H1N1 (1); osteomyelitis (1); sepsis (1); sinusitis (2); staphylococcal infection (3); upper respiratory tract infection (1); urinary tract infection (1)	[redacted]	Bacterial sepsis (1); bronchopneumonia (1); catheter related infection (1); Escherichia sepsis (1)
Injury, poisoning and procedural complications	[redacted]	Accidental overdose (1); incision site haemorrhage (1); infusion related reaction (1); overdose (5); post lumbar puncture syndrome (1); spinal fracture (1); subdural haemorrhage (1); thermal burn (1)	[redacted]	Medical device complication (1); thrombosis in device (1)
Investigations	[redacted]	Alanine aminotransferase increased (2); aspartate aminotransferase increased (2); blood bilirubin increased (1); body temperature increased (1); c-reactive protein increased (4); hepatic enzyme increased (1); liver function test abnormal (1); prothrombin time prolonged (1)	[redacted]	NR
Neoplasms benign, malignant and unspecified (incl cysts and polyps)	[redacted]	Kaposi's sarcoma (1); leukaemia (1)	[redacted]	NR
Nervous system disorders	[redacted]	Aphasia (6); ataxia (2); cognitive disorder (1); dysarthria (1); encephalopathy (6); generalised tonic-clonic seizure (1); headache (2); intention tremor (1); leukoencephalopathy (1); motor dysfunction (1); paraesthesia (1); seizure (3); tremor (8)	[redacted]	Convulsion (1); epilepsy (1); somnolence (1); syncope (1)
Psychiatric disorders	[redacted]	Agitation (1); confusional state (1); disorientation (1)	[redacted]	NR
Skin and subcutaneous tissue disorders	[redacted]	Dermatitis contact (1); rash maculo-papular (1)	[redacted]	NR
Vascular disorders	[redacted]	Hypotension (1); thrombosis (1); vena cava thrombosis (1)	[redacted]	Hypertension (1)

NR - not reported

Table 15: Comparison of SAEs in MRD+BCP-ALL patients (Pooled BLAST + MT103-202) with known safety profile from relapsed and refractory Ph-BCP-ALL patients (Pooled MT103-206 + MT103-211 + TOWER)

Event	MRD+BCP-ALL patients (Pooled BLAST + MT103-202), n=137	Ph-BCP-ALL patients (Pooled MT103-206 + MT103-211 + TOWER), n=NR*
Treatment-emergent grade ≥3 AEs	[redacted]	[redacted]
Treatment-emergent serious AEs	[redacted]	[redacted]
Treatment-related AEs	[redacted]	[redacted]
Treatment-related serious AEs	[redacted]	[redacted]
Treatment-emergent EOIs grade ≥3	[redacted]	[redacted]
Treatment-emergent EOIs grade ≥4	[redacted]	[redacted]
EOI grade ≥3	[redacted]	[redacted]
EOI grade ≥4	[redacted]	[redacted]
Neurological AEs	[redacted]	[redacted]
Cytokine release syndrome	[redacted]	[redacted]
Medication errors	[redacted]	[redacted]

* *Total pooled n not reported. Individual studies reported as n=36 (MT103-206), n=189 (MT103-211), n=NR (TOWER)* [43]

4.3 Study included as comparator

Study 20120148 was a retrospective study that collected data on PH-BCP-ALL patients who were in complete haematological remission with MRD (see Table 16). The rationale for the study was to provide a frame of reference from which to compare the single-arm BLAST study of blinatumomab.[33] Treatment and outcome data were collected retrospectively from study groups across Europe and Russia (see CS,[1] Section B.2.9). MRD assessment was by PCR or by flow cytometry at a reference lab.[33] Study 20120148 collected OS and RFS data, but did not provide data on AEs. Within the limitations of the study design, the study was well conducted; as noted in Section 4.1.4, single-arm and retrospective studies are associated with known biases.

Eligibility criteria for Study 20120148 were available from the CS[1] (Section B.2.9) and the US NIH clinical trials registry;[41] these are presented below.

Study 20120148 inclusion criteria

Patients with Ph- BCP-ALL with haematological CR (defined as less than 5% blasts in bone marrow after at least 3 intensive chemotherapy blocks, and who met the following criteria:

- Detection of MRD (molecular failure or molecular relapse) at a level of $\geq 10^{-4}$ by PCR or $\geq 10^{-3}$ by flow cytometry at a reference lab

- Age 15+ at time of initial diagnosis of ALL. For patients 15-17 years of age at diagnosis, patients were not allowed to be enrolled in a paediatric trial, i.e. had to be treated according to adult protocols
- Initial diagnosis of ALL in the year 2000 or later
- History of ALL treatment (including response to first therapy, number of prior relapses) is available
- Relapse status and disease follow-up after time point of MRD detection is available.

Study 20120148 exclusion criteria

- Patients with extramedullary disease at timepoint of MRD detection
- Use of blinatumomab within 18 months of MRD detection
- Allogeneic HSCT prior to MRD detection at required level.

From the data collected in Study 20120148, a direct comparison analysis set (DCAS) was selected to act as matched controls for the BLAST study (see CS,[1] Section B.2.9). Additional criteria were applied in order to produce the DCAS. Data from Russian patients were excluded because MRD levels were not quantified. Patients were in their first haematological remission (CR1) only. Only patients aged 18 years or older at the MRD baseline date were included. Time to relapse had to be greater than 14 days from the date of MRD detection.

Baseline characteristics

Baseline characteristics for patients in Study 20120148 are presented in Table 17. Most of the patients in the DCAS were from [redacted] (see CS,[1] Section B.2.9.3). The DCAS included [redacted] from the UK.[33]

Table 16: Characteristics of retrospective control study

Study	Reference(s)	Study design	Population	Number of patients	Intervention	Date of initial diagnosis	Outcomes
Study 20120148 NCT02010931	Amgen Inc. Observational Research Study Report 2017[33]	Retrospective, international, multicentre	Adult Ph-BCP-ALL patients in haematological CR with MRD	Data collected for 287 patients [redacted] patients selected for DCAS	Standard care chemotherapy regimens, according to national treatment or study group protocols	2000 - 2014	Haematological RFS rate, OS, Mortality rate 100-days following HSCT

Information from CS Sections B.2.2 and B.2.9, CS clarification response[17] question A13 and Amgen Study Report[33]
DCAS= direct comparison analysis set

Table 17: Baseline characteristics of Study 20120148 direct comparison analysis set

Demographic	Study 20120148 [redacted]* Prior to adjustment**
Male sex, n (%)	[redacted]
Median age (range), years	[redacted]
Age, n (%) ≥18 to <35 years	[redacted]
≥35 to <55 years	[redacted]
≥55 to <65 years	[redacted]
≥65 years	[redacted]
Relapse history, n (%) First CR	[redacted]
Second CR	[redacted]
Third CR	[redacted]
Baseline MRD levels, n (%) $\geq 10^{-1}$ <1	[redacted]
$\geq 10^{-2}$ $<10^{-1}$	[redacted]
$\geq 10^{-3}$ $<10^{-2}$	[redacted]
$<10^{-3}$	[redacted]
Philadelphia chromosome disease status Negative	[redacted]
Confirmed t(4;11) Translocation / MLL-AF4+	[redacted]
Time from diagnosis to baseline (months) mean (SD)***	[redacted]
WBC at diagnosis (≥30,000/mm^3)	[redacted]

Adapted from CS Section B.2.9 Table 28 and Appendix L Table 86 and Amgen Study report [44]
**Patients ≥18 years old with MRD load ≥1 × 10−3 detected by FC or PCR in CR1, time to haematological relapse >14 days after MRD diagnosis. **For details on adjustment see ERG report Section 4.4.*
****Time from initial diagnosis to baseline MRD status defined as the earliest MRD detection date following complete remission after at least three blocks of chemotherapy[33]*
CR: complete remission; DCAS: direct comparison analysis set. WBC: white blood cell

4.4 Indirect comparison

Owing to the lack of randomised data to inform the comparative effectiveness of blinatumomab versus standard care chemotherapy, the company performed an analysis based on the historical cohort DCAS, designed *post hoc* to include patients resembling those enrolled into the BLAST study. RFS and OS outcomes were considered; other outcomes listed in the final NICE scope[22] were not reported for the indirect comparison. Propensity score methods based on inverse probability of treatment weighting (IPTW) were used.

The data used to inform the analysis are described in Section 4.4.1. The methods used to estimate treatment effectiveness are described in Section 4.4.2 and are subsequently critiqued according to the items in the Quality of Effectiveness Estimates from Non-randomised Studies (QuEENS) checklist.[45]

4.4.1 Critique of included studies

The effectiveness of blinatumomab was informed by the BLAST study (n=[redacted]), as summarised in Section 4.2, whilst the effectiveness of standard care was informed by the historical comparator DCAS (n=[redacted]), as summarised in Section 4.3. Baseline characteristics of the full BLAST study and historical control DCAS are compared in Table 28 of the CS. As noted by the company, there were key differences between the two populations in terms of Ph status and relapse history.[1]

The BLAST primary analysis set (PAS) was trimmed to overlap with the historical comparator DCAS. The two key criteria defining this subgroup are the restriction to CR1 and Ph- individuals only; the full criteria are listed below:

- Ph- BCP- ALL;
- First complete haematological remission (CR1);
- MRD+ at a level of >1 x 10-3;
- ≥18 years old at MRD positivity (historical control study [Study 20120148]) or first blinatumomab treatment (BLAST [Study MT103-203]);
- Complete baseline covariate set;
- Time to relapse greater than 14 days from MRD detection (applied to historical control study data).

Trimming resulted in a subgroup of [redacted] patients for the BLAST PAS.

The timing of MRD assessment following diagnosis also varied between the BLAST PAS and historical comparator DCAS, and within different study groups contributing to the historical comparator DCAS. In order to align the baseline dates and to reduce bias due to the definition of MRD baseline date, patients in the historical comparator DCAS were excluded if their time to relapse was less than 14 days (the median time between MRD detection and first blinatumomab dose for BLAST patients). The baseline date for patients within the historical comparator study was set equal to their MRD detection date plus 14 days. This led to the exclusion of four patients from the historical control study, due to relapse during the first 14 days after MRD baseline (see company's clarification response,[17] question A11).

The cases from the control study were recruited from the year 2000 onwards, as opposed to the BLAST study, in which cases were recruited from 2010 onwards. There have been some changes to induction treatment, which may mean that more recently treated patients have lower rates of MRD positivity and lower rates of relapse. However, there is an absence of evidence for this in UK-treated patients. Clinical

advice received by the ERG suggests that it is broadly reasonable to assume that treatments received by patients from 2000 onwards would be similar to current practice.

4.4.2. Critique of methods for estimating comparative effectiveness

Description of analysis performed by company

Differences between the BLAST PAS and historical comparator DCAS with respect to key baseline characteristics (prior to propensity score adjustment) are shown in Table 18. Balance with respect to individual covariates was assessed by the company using two methods: (i) univariate regression models were constructed to investigate the association between the treatment group (as the predictor), on each baseline characteristic (as the outcome variable) individually, using linear and logistic regression for continuous and binary baseline characteristics respectively, with results reported as *p*-values, and (ii) standardised mean differences between the two groups were calculated (formulae presented in CS Appendix L). The CS states that the criteria for concluding that adequate balance was achieved were: (i) non-significant *p*-values and (ii) standardised differences less than 0.2, with "best balance" achieved with standardised differences less than 0.1.[1]

Before applying the propensity score weighting, four of the listed covariates had *p*-values which were less than 0.05: age; country; time from diagnosis to baseline (months), and prior chemotherapy. The absolute standardised differences ranged from 0 to 0.56, with standardised differences greater than 0.2 observed for WBC at diagnosis (continuous) in addition to the four covariates listed above. Only two covariates (gender and T411mll4 mutation) exhibited standardised difference less than 0.1, which is indicative of good balance between the groups.

Due to the observed differences in baseline characteristics between the BLAST PAS and the historical control DCAS, IPTW based on a propensity score model was used. The overall aim of the procedure is to create balance between the two groups by producing a weighted sample that mimics the effect of randomisation in an RCT.[45] The propensity score model estimates the probability of being assigned to the treatment group as a function of a set of observable covariates. These propensity scores are used to construct weights that are applied to the observed data. Several weighting schemes may be considered, each of which results in different interpretations of the resulting treatment effect. The average treatment effect (ATE) measures the expected gain from the treatment for a randomly selected individual (across both samples) and is most appropriate when the treatment is relevant to the entire population represented by the data. Weights are applied to both the BLAST PAS and the historical control DCAS patients (see CS,[1] Appendix L). The average treatment effect on the treated (ATT) is relevant when the interest lies on the effect of treatment only for those who are treated (rather than the population of both treated and untreated patients). No weighting is applied to the blinatumomab patients, whilst patients in the historical control arm are weighted to match those in the treated study.

The company used ATT weights to inform the health economic model (see Section 5). The justification for this was that results based on ATT weights can be generalised to the population of patients from BLAST, which represents the prospectively selected anticipated licensed population, rather than the combined population of the BLAST study and the historical control study. Results using ATE weights are presented in CS Appendix L, and were used for a sensitivity analysis. In order to adjust for potential instability caused by very large weights, stabilised weights (applied to both the ATT and ATE analyses) were presented by the company, whereby the weight is multiplied by the marginal probability of receiving the actual treatment received.[46] This results in a smaller effective sample size of [redacted] (see company's clarification response,[17] A8 Additional Query). The stabilised weights were used to produce the estimates of treatment effect presented in the clinical effectiveness section of the CS. The company acknowledged that there was a lack of consistency between the results presented in the clinical effectiveness section and those used to inform the health economic model, but stated that they are confident that the application of the stabilised ATT (sATT) weights to the cost-effectiveness analyses would have "*no impact*" (see company's clarification response,[17] question A11).

Candidate variables for the company's propensity score model were chosen through discussion amongst the study team and clinicians. As stated in the company's response to clarification[17] (question A8), the majority of covariates were chosen based on prognostic factors that have been identified for ALL in published literature and to account for potential regional differences in treatment practices. Candidate variables included: age at primary diagnosis; sex; country; presence and type of an cytogenetic and molecular aberrations; time from primary diagnosis to MRD baseline data (months); baseline MRD level (ordinal variable, treated as continuous in the model); WBCs at diagnosis, and type of prior chemotherapy (binary: GMALL, other). The final propensity score model was chosen by including all candidate variables and two-way interactions into a logistic regression with treatment as the binary response. A stepwise selection algorithm was used with inclusion into the final model based on statistical significance ($p<0.30$).

Table 18: Covariate balance between BLAST PAS and historical control study, before and after adjustment using ATT weights (reproduced from company's clarification response question A8)

Characteristic	Unweighted				IPTW			
Mean (SD)/n (%)	Control	Blinatumomab	Standard Difference	*p*-value	Control	Blinatumomab	Standard Difference	*p*-value
Age at primary diagnosis (years)	[redacted]							
Gender (female)								
Country (not Germany)								
MRD at Baseline (recoded)								
Time from diagnosis to baseline (months)								
WBC at diagnosis (>30,000/mm^3)								
WBC at diagnosis (continuous, log10)								
T411mll4 mutation (Yes)								
Prior chemotherapy (GMALL)								

Table 19: Summary of propensity score model covariates (modified from company's propensity score analysis report)

Covariate	estimate (SE)	*p*-value
age at primary diagnosis (years)	[redacted]	[redacted]
time from diagnosis to baseline (months)	[redacted]	[redacted]
MRD level at baseline	[redacted]	[redacted]
type of prior chemotherapy	[redacted]	[redacted]
Not GMALL	[redacted]	[redacted]
time from diagnosis to baseline (months) x type of prior chemotherapy	[redacted]	[redacted]
Not GMALL	[redacted]	[redacted]

SE - standard error
p-value from Wald Chi-Square statistic

Balance diagnostics after applying ATT weights to the historical control DCAS are presented in Table 18, based on the company's clarification response[17] (question A8). After applying ATT IPTW weights, [redacted] Balance diagnostics for the sATT weights used to estimate the treatment effects are shown in the CS[1] (Table 86, page 220). [redacted]

The derived sATT propensity score weights were used to perform a weighted Cox proportional hazards analysis and therefore estimate the hazard ratio (HR), providing a treatment effect comparing blinatumomab to standard care. Analyses were conducted separately for both RFS and OS. The primary analysis considered just one covariate (allowing a treatment effect), and an additional analysis was conducted including a time-dependent covariate for HSCT to account for differences between transplant rates observed between BLAST and the historical cohort. Analyses were conducted in SAS. The adjusted Kaplan-Meier plots using ATT weights presented by the company are shown in ***[redacted]4 and ***[redacted]5 for RFS and OS, respectively; estimated treatment effects are summarised in Table 20.

[redacted]

[redacted] After applying sATT weights, the 18-month OS probability with standard care chemotherapy, without censoring for HSCT, was [redacted] (CS Appendix L) and the median OS was slightly longer than prior to weighting, at [redacted]. For the BLAST PAS with sATT weights [redacted] without censoring for HSCT, the 18-month OS probability was [redacted]). Median OS for the BLAST patients [redacted].

[redacted] After applying sATT weights, the 18-month RFS probability with standard care chemotherapy, without censoring for HSCT, was [redacted] (CS Appendix L) and the median RFS was [redacted]). For the BLAST PAS [redacted] without censoring for HSCT, the [redacted] RFS probability was [redacted]). Median RFS for the BLAST patients was [redacted].

Table 20: Estimated treatment effects based on sATT weights

Outcome	Median (months)		HR (95% CI)	
	Standard care [redacted]	Blinatumomab [redacted]	primary analysis	covariate for HSCT
RFS	[redacted]	[redacted]	[redacted]	[redacted]
OS	[redacted]	[redacted]	[redacted]	[redacted]

[redacted]

[redacted]

Note that after application of the sATT weights, the effective sample sizes are

[redacted]

[redacted]

Note that after application of the sATT weights, the effective sample sizes are

[redacted]

Critique of the analysis

The ERG considers that the IPTW method used by the company is appropriate and that other methods are not suitable in this case due to the limited sample size. The method makes two important assumptions. Firstly, the methods assume that there is *no unobserved confounding* (also described as selection on observables). When estimating treatment effects based on non-randomised data it is possible that a patient received a particular treatment because of some (observable or unobservable) factors. Unless properly accounted for, this will lead to selection bias in the estimated treatment effect.[45] Selection on observables implies that all factors which determine treatment and are correlated with the outcome are observable, and hence can be accounted for in the propensity score model. There may be unobservable factors which determine treatment allocation, but these are not correlated with the outcome. Secondly, the *overlap assumption* is also required. This means that, for any combination of covariates, it is possible for individuals to be allocated to either the treatment or control group, ruling out the possibility that individuals with certain observable characteristics are always in one group and

never in the other.[45] Weaker versions of both of these assumptions are required for the validity of the ATT weights, compared with the requirements of the ATE weights.

The analysis was based on a subset of individuals, the BLAST PAS, rather than the whole study population. This "trimming" is generally required in order to meet the overlap assumption when the initial overlap between the two populations is poor.[45] However, this redefines the interpretation of the estimated treatment effects. The ATT weights presented as the company's primary analysis represent the average treatment effect for the population of the BLAST PAS (n=[redacted]), which was chosen to overlap with the historical control study, rather than the full BLAST study population. The company's justification of the choice of ATT weights (rather that the ATE weights that were pre-specified in the protocol) due to the BLAST study being in line with the anticipated marketing authorisation is therefore not consistent with the interpretation of the resulting estimates, which are representative of the subpopulation only.

The assumptions required by each of the weighting methods are described in the CS, however, it is not clearly stated whether there is reason to believe that the stronger assumptions required for the validity of the ATE weights may not be met. ATT weights were used in the company's health economic base case analysis despite the fact that there was "*less improvement in covariate balance after weighting when using ATT*".[44] Overall, the ERG does not consider that the company's choice of weights for the base case analysis has been clearly justified. There was also a lack of clarity and consistency caused by the use of sATT weights to estimate treatment effects, and the application of standard ATT weights in the cost-effectiveness analysis. The ERG considers that that the use of the standard ATT weights was appropriate.

Clinical advisors to the ERG considered that the candidate variables considered by the company were generally appropriate; however, they drew attention to the potential for unobserved confounders related to HSCT status. As noted within the CS[1] (Section B.2.9.5, page 75), transplanted patients may be systematically different in terms of both measured and unmeasured factors (such as availability of a suitable donor). The HSCT rate is higher in the BLAST study ([redacted]) than the historical control study ([redacted]), and the CS states that the comparison is vulnerable to HSCT being a confounding factor.

The ERG believes that the choice of a logistic regression model was appropriate. However, the inclusion of covariates in the final model was based on statistical significance only. The CS does not present any checks (e.g. model diagnostic plots) for the final model. After applying ATT weights to the historical control DCAS, the company's pre-specified criteria for judging balance between the populations was met. This was not true for the sATT weights used to estimate treatment effects, as three covariates (age,

time from diagnosis to baseline, WBC at diagnosis) still had standardised differences greater than 0.2. However these results were not used to inform the cost-effectiveness analysis.

Furthermore, it should be emphasised that the propensity score weights (hence, also the adjusted Kaplan-Meier survival curves) are estimates with associated measures of uncertainty (e.g. SEs). It is unclear (although unlikely) that this has been accounted for in the estimation of treatment effects, hence the reported confidence intervals of the treatment effects are likely to underestimate the associated uncertainty. The ERG therefore considers that the reported treatment effects should be interpreted with caution.

4.5 Additional work on clinical effectiveness undertaken by the ERG

No additional work on clinical effectiveness was undertaken by the ERG.

4.6 Conclusions of the clinical effectiveness section

4.6.1 Completeness of the CS with regard to relevant clinical studies and relevant data within those studies

The clinical evidence presented in the CS is based on a systematic review of adult BCP-ALL patients with MRD positivity after treatment. The company's study selection eligibility criteria were consistent with the decision problem outlined in the final NICE scope,[22] except that the comparator "monitor for relapse" was not included; the ERG's clinical advisors noted that some older and less fit patients may not be able to receive HSCT or to tolerate chemotherapy, but may be able to tolerate blinatumomab. It is unclear whether any relevant comparator data exist within this subgroup. Overall, the ERG believes that whilst the searches conducted by the company were flawed, it is unlikely that any relevant studies of blinatumomab in adult BCP-ALL patients with MRD positivity after treatment have been missed.

Evidence of the effectiveness of blinatumomab was provided from two single-arm open-label studies, BLAST (n=116) and MT103-202 (n=20), with no internal control group against which to estimate a treatment effect. Comparator data relating to standard care chemotherapy were provided from one historical control study, Study 20120148 (n=287), that analysed data from existing clinical databases.

AE data for blinatumomab were presented for BLAST and MT103-202. There were no data on AEs or HRQoL from historical control study 20120148.

4.6.2 Summary of clinical effectiveness outcomes reported in the CS in relation to relevant population, interventions, comparator and outcomes

Clinical advice received by the ERG suggested that baseline demographics and prior treatment in the BLAST study were broadly generalisable to the population of MRD+BCP-ALL patients in England.

The ERG notes that there will be a small population of patients who are unable to undergo HSCT or to tolerate chemotherapy who are unlikely to be represented within the BLAST population.

From the 116 patients in BLAST, median OS was [redacted], with an OS at 18 months follow-up of [redacted]. From 110 patients providing RFS data from BLAST, median RFS was [redacted]; RFS at 18 months was [redacted]. BLAST measured HRQoL using the EORTC QLQ-C30 and the EQ-5D. Based on the EORTC QLQ-C30, patients reported [redacted]. By the end of the core study, [redacted] HRQoL as measured by EQ-5D did not change significantly from baseline to the end of the core study. [redacted] participants experienced at least one treatment-emergent AE. Events occurring in ≥20% of participants included: [redacted]. The most common EOIs of blinatumomab were: neurological events [redacted].

Comparative effectiveness was estimated by applying sATT propensity score weights to the standard care chemotherapy arm. Due to differences between the populations of BLAST and the historical control study, this was based on a subset of the original study populations which were restricted to Ph- and CR1 individuals only (BLAST PAS n=[redacted], historical control DCAS n=[redacted]). The resulting treatment effect estimates therefore reflect a narrower population than that defined in the final NICE scope[22] and the wording of the anticipated marketing authorisation.[24]

For the BLAST PAS, the 18-month OS probability was [redacted].

For the BLAST PAS, the 18-month RFS probability was [redacted]

4.6.3 Uncertainties surrounding the reliability of the clinical effectiveness evidence

A key limitation of the effectiveness evidence is the design of the included studies. The two blinatumomab studies were well conducted, however single-arm studies are subject to several biases.[34] Comparative effectiveness was estimated using propensity score methods which the ERG considers to have been appropriately applied by the company; however, the estimation of treatment effects based on non-randomised data is still subject to limitations, namely that it is not possible to account for unobserved confounders, and the company states that the comparison is vulnerable to HSCT being a confounding factor.

Treatment effects (HR) appear to have been calculated ignoring the uncertainty associated with the estimated propensity score weights, and therefore it is likely that the estimates presented within the CS underestimate the total uncertainty of the reported HR, resulting in erroneously narrow confidence intervals. The ERG therefore considers that the reported treatment effects should be interpreted with caution.

A further limitation of the available evidence relates to generalisability to the full population outlined in the final NICE scope and the anticipated license.[22] On the basis of clinical advice, the ERG considers the population characteristics of the BLAST PAS and the historical control DCAS to be representative of Ph- CR1 patients with MRD+ BCP-ALL. However, there is no evidence to inform the comparative effectiveness of blinatumomab compared with standard care chemotherapy in patients with CR2+ and/or Ph+ disease. In addition, no evidence is reported for blinatumomab versus monitoring for patients who are unable to receive HSCT or to tolerate chemotherapy but who would be able to tolerate blinatumomab.

5. COST EFFECTIVENESS

This chapter presents a summary and critical appraisal of the methods and results of the company's review of published economic evaluations and the *de novo* health economic analysis presented within the CS.[1] All analyses presented in this chapter including the Patient Access Scheme for blinatumomab.

5.1 ERG comment on the company's systematic review of cost-effectiveness evidence

5.1.1 Description of company's systematic review of cost-effectiveness evidence

A single search strategy reported in Appendix C (also used in the identification, selection and synthesis of clinical evidence) of the CS was used to identify the following study types: (i) economic analyses of all interventional therapies for adult ALL patients with MRD; (ii) HRQoL studies in patients with MRD+ BCP-ALL, and (iii) studies assessing the economic burden of patients with MRD+ BCP-ALL. The search strategies in both the database and website searches were fully reported. The records retrieved from the search were for all MRD+ ALL patients.

The following sources were searched: MEDLINE in Process [via PubMed], EMBASE [host not reported], Cochrane Database of Systematic Review [via Wiley Online Library], Cochrane Central Register of Controlled Trials [via Wiley Online Library], Database of Abstracts of Reviews of Effectiveness [via CRD], NHS Economic Evaluation Database [via CRD] and the Health Technology Assessment database [via CRD].

The ERG's concerns regarding the limitations of the company restrictions applied to the search strategy (MRD terms, study design and language limits) have been previously described in Section 4.1.1. Following the consultation with clinical experts, the ERG considers that the search is sufficiently comprehensive to retrieve important citations relating to all eligible studies.

The company's inclusion and exclusion criteria are reproduced in Table 21. The company's review included adult ALL patients with MRD-positivity after treatment and was not restricted by intervention. However, the company's searches did not identify any existing economic evidence relating to adult ALL patients with MRD-positivity after treatment.

Table 21: Company's review of existing economic studies - inclusion and exclusion criteria (adapted from CS, Appendix G)

	Inclusion criteria	Exclusion criteria
Population	Adult ALL patients with MRD-positivity after treatment	Paediatric patients MRD- ALL patients
Intervention/ comparator	Any interventional therapies	None
Outcomes	Cost effectiveness Measures of cost effectiveness (e.g. cost per QALY gained)	Non-economic outcomes
Study design	Economic analyses and HTA reports	Non-economic study designs

5.2 Description of the company's model

5.2.1 Model scope

As part of its submission to NICE, the company submitted a fully executable health economic model programmed in Microsoft Excel®. The scope of the company's health economic analysis is summarised in Table 22. The company's model assesses the cost-effectiveness of blinatumomab versus standard care chemotherapy in adult patients with Ph- MRD+ BCP-ALL in CR1. The incremental health gains, costs and cost-effectiveness of blinatumomab versus standard care are evaluated over a 50-year time horizon from the perspective of the UK NHS and Personal Social Services (PSS). Cost-effectiveness is expressed in terms of the incremental cost per quality-adjusted life year (QALY) gained. All costs and health outcomes are discounted at a rate of 3.5% per annum. Unit costs are valued at 2015/16 prices.

Table 22: Summary of company's health economic model scope

Population	Patients with Ph- MRD+ BCP-ALL in CR1
Intervention	Blinatumomab (up to 4 cycles)*
Comparator	Standard care - chemotherapy regimen assumed be comprised of vincristine, prednisolone, mercaptopurine, methotrexate and prophylaxis against CNS relapse using intrathecal methotrexate (treatment up to 2 years)
Primary health economic outcome	Incremental cost per QALY gained
Perspective	NHS and PSS
Time horizon	50 years
Discount rate	3.5% per annum
Price year	2015/2016

** All patients receiving blinatumomab are also assumed to receive prophylaxis against CNS relapse*
NHS – National Health Service; PSS – Personal Social Services

Population

The population considered within the company's health economic model is defined according to the characteristics of patients enrolled into the BLAST study and the historical comparator study who met the criteria stated in Section 4.4.1. These subgroups of the full study populations are described as the historical comparator DCAS ([redacted]) and the BLAST PAS ([redacted]). The company's health economic analysis is based on ATT propensity weights, rather than the ATE weights that are presented in CS

Appendix L. This approach was taken on the basis that the analysis based on the ATT weights *"can be generalised to the population of patients in BLAST rather than the combined populations of the BLAST and historical control studies"* (see Section 4.4).

It should be noted that the company's health economic analysis reflects a population of patients who are likely to be able to tolerate chemotherapy; clinical advisors to the ERG noted that owing to its toxicity profile, blinatumomab may be a treatment option for patients who are not fit enough to receive HSCT or to tolerate cytotoxic therapy; this subgroup is unlikely to be reflected by the population captured within the company's model. In addition, the company's economic analysis excludes two further subgroups of patients who were included in the BLAST study: (i) patients with Ph- MRD+ BCP-ALL who are in second or subsequent haematological remission (CR2+), and (ii) patients with Ph+ MRD+ BCP-ALL. The population considered within the model is therefore narrower than the anticipated marketing authorisation for blinatumomab (treatment of adults with MRD+ BCP-ALL).[24] The CS states that undertaking a formal economic analysis of blinatumomab in the broader patient population, which also includes patients in CR2+, was infeasible due to a lack of comparator data. However, despite the absence of any clinical or economic evidence to support the analysis of blinatumomab in these missing subgroups, the CS states *"due to the substantial unmet need across all sub-populations blinatumomab should be considered for use in alignment with its full anticipated marketing authorisation"* (CS,[1] page 15). The CS also states that it anticipates that blinatumomab would be used as early as possible in the treatment pathway. These issues are discussed further in Section 5.3.

Intervention

In the BLAST study, blinatumomab was administered as a continuous IV infusion at a dose of 15µg/m^2 per day for 4 weeks, followed by a 2-week treatment-free period. Patients could receive up to four consecutive treatment cycles of blinatumomab. In contrast, the model assumes that a single cycle of blinatumomab treatment is comprised of a continuous IV infusion at a dose of 28µg/day for 28 days, followed by a 14-day treatment-free interval. This is in line with the anticipated marketing authorisation for blinatumomab.[24] The model assumes that patients receive 1 cycle of induction treatment followed by up to 3 additional cycles of consolidation treatment.

Comparator

The comparator included in the company's model is standard care chemotherapy. Health outcomes for the comparator group are based on the historical control DCAS, whilst the costs of standard care are modelled according to the maintenance chemotherapy regimen for non-transplant patients used in the UKALL14 trial:[19]

- Vincristine 1.4mg/m^2 (maximum 2mg/dose) IV every 3 months for up to 2 years

- Prednisolone 60mg/m^2 orally 5 days every 3 months for up to 2 years
- Mercaptopurine 75mg/m^2 orally daily for up to 2 years
- Methotrexate 20mg/m^2 orally once weekly for up to 2 years
- Prophylaxis against CNS relapse using intrathecal methotrexate 12.5mg every 3 months for up to 2 years.[1]

The final NICE scope[22] also included a further comparator of "monitor for relapse" (no active treatment); the CS[1] justifies the exclusion of this comparator by stating: *"Based on expert clinical opinion it is highly unlikely that MRD+ patients who have a high risk of relapse would solely be monitored for relapse without any treatment"* (CS,[1] page 15). However, clinical advisors to the ERG noted that this comparator is relevant for those patients who are unable to undergo HSCT or to tolerate chemotherapy, but are able to tolerate blinatumomab. This comparator therefore should have been explored in the company's economic analysis.

5.2.2 Description of the company's health economic model structure and logic

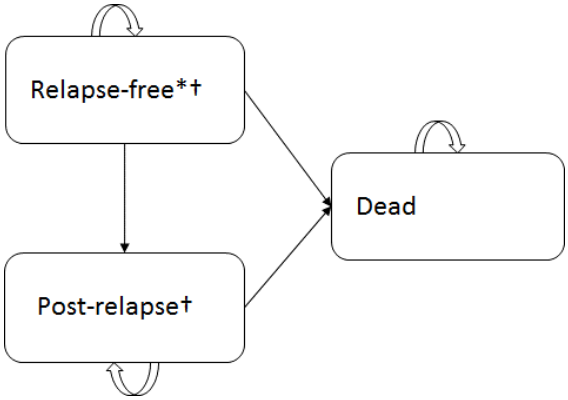
The company's model is comprised of a main structure which reflects RFS and OS outcomes from the BLAST PAS and historical control study DCAS, as well as two linked sub-models which estimate additional costs and HRQoL decrements associated with HSCT given before and/or after relapse. The subsequent sections describe the main model structure and the two HSCT sub-models separately.

Main partitioned survival model structure

The company's model adopts a partitioned survival approach based on three health states: (1) relapse-free; (2) post-relapse, and (3) dead (see Figure 6). Patients enter the model in the relapse-free state with an initial age of 45.38 years. Health state transitions are estimated over a total of 2,607 weekly cycles (approximately 50 years); at this timepoint, more than 99.9% of patients in each treatment group have died. The probability of being alive and relapse-free at any time *t* is based on a parametric (Gompertz) model fitted to the treatment-specific RFS time-to-event data from the BLAST PAS and the historical control DCAS with ATT weights. The probability of being alive at any time *t* is modelled using a parametric (log normal) mixture cure model fitted to the OS time-to-event data from the BLAST PAS and the historical control DCAS with ATT weights, as well as a separately estimated general population survivor function. The latter OS survivor function is estimated using age- and sex-specific mortality risks from life tables which are uplifted by a factor of 4 (based on the NICE Appraisal Committee's preferred assumption within the appraisal of inotuzumab ozogamicin for treating R/R ALL[47]) to reflect the potential long-term effects of complications of cytotoxic chemotherapy, and/or allogeneic HSCT on survival. Within the model trace, the probability of surviving during each model cycle is determined by the cumulative OS probability at the end of the previous model cycle and the maximum OS hazard for

the current cycle derived from the parametric OS cure model and the uplifted general population survival curve. The probability of being alive and in the post-relapse state at any time *t* is calculated as the difference between the cumulative survival probabilities for OS and RFS.

Figure 6: Company's model structure



** RFS time divided into time on treatment and post-discontinuation*
† Patients may enter state-specific HSCT sub-model

Pre-relapse and post-relapse HSCT sub-models

Given the use of a partitioned survival approach in which health states are defined according to patients' survival and relapse status, the company's model structure does not explicitly account for differential RFS and OS impacts for those patients who receive HSCT; within the model, the proportionate use of pre-relapse HSCT is causally unrelated to RFS and OS events, whilst post-relapse HSCT use is partially dependent on RFS. In both treatment groups, the probability that a patient undergoes HSCT is approximated using separate pre-relapse and post-relapse HSCT sub-models in order to attribute costs and QALY losses associated with this intervention.

For patients who are relapse-free, the modelled (time-invariant) 6-monthly probability of receiving HSCT was calibrated such that the predicted cumulative probability of having undergone pre-relapse HSCT at 48 months matches the observed probability from the BLAST study and the historical control study. Beyond 48 months (based on the time of the last observed pre-relapse HSCT in BLAST and the historical control study), the model assumes that patients in the relapse-free health state of the main partitioned survival model cannot subsequently undergo HSCT, unless they relapse and enter into the post-relapse HSCT sub-model. Whilst the modelled proportion of patients receiving pre-relapse HSCT is dependent on the RFS function, OS in the main partitioned survival model is unaffected by the pre-relapse HSCT sub-model. After undergoing HSCT, 6-monthly follow-up costs and QALY losses are

estimated using HSCT follow-up data from NHS Blood and Transplant[48] up to 2-years, and using uplifted general population survival rates thereafter.

With respect to the post-relapse HSCT sub-model, the per cycle probability of receiving HSCT after relapse is calculated within the model by determining the number of RFS events since the end of the previous 6-month HSCT sub-model cycle (derived from the modelled RFS curve) and a time-invariant treatment group-specific probability that an RFS event is death. The 6-month probability of undergoing post-relapse HSCT is determined by two factors: (i) the probability of undergoing HSCT for those patients who have not previously undergone HSCT whilst relapse-free, and (ii) the probability of undergoing HSCT for those patients who have previously undergone HSCT whilst relapse-free. As the model structure does not capture a patient's history of HSCT in the pre-relapse state, the model necessarily employs an assumption which attempts to estimate the probability of receiving post-relapse HSCT according to whether patients have undergone pre-relapse HSCT or not. In simple terms, the model is intended to assume that patients with pre-relapse HSCT do not relapse until all patients without pre-relapse HSCT have relapsed (see company's clarification response,[17] question B32, although the ERG notes that the implementation actually requires further assumptions about when the HSCT probability switches). As with the pre-relapse HSCT sub-model, 6-monthly follow-up costs and QALY losses are estimated using HSCT follow-up data from NHS Blood and Transplant[48] up to 2-years, and using uplifted general population survival rates thereafter.

Modelling HRQoL impacts

The model assumes that HRQoL is principally determined by relapse status, time spent alive and relapse-free and treatment received. Within the blinatumomab group, the model applies different health utilities in the relapse-free state over time; HRQoL is also assumed to differ for patients who are still receiving treatment and for those who have discontinued blinatumomab. Within the standard care group, the model applies fixed utilities for the relapse-free and relapsed states for up to 5-years. Within both treatment groups, HRQoL in the relapse-free state beyond 5-years is assumed to reflect that of the age- and sex- adjusted general population, less a constant utility decrement of 0.02, which is assumed to reflect long-term impacts associated with radiotherapy, chemotherapy, and HSCT. In addition, further time-dependent QALY losses are applied for those patients undergoing HSCT for up to 5 years. A further QALY loss is also applied to account for patients' proximity to death.

Modelled treatment pathway and associated costs

The company's model includes the following cost components: (i) drug acquisition; (ii) drug administration; (iii) health state resource use; (iv) HSCT; (v) salvage therapy costs, and (vi) a cost associated with death.

Within the blinatumomab group, the model assumes the following treatment pathway:

- Patients receive up to four cycles of blinatumomab irrespective of relapse status (experiencing relapse does not trigger the discontinuation of blinatumomab). Each cycle is comprised of 28 days receiving 28μg blinatumomab followed by 14 days without treatment. The model calculates blinatumomab costs based on the unweighted mean proportion of those patients starting the cycle and those still on treatment at the end of the cycle.
- Prophylaxis against CNS relapse is given to all patients in the relapse-free health state for up to 2 years, unless they progress to pre-relapse HSCT or die. The prophylaxis regimen is comprised of 15mg methotrexate, 40mg cytarabine and 4mg dexamethasone given once every 13 weeks. All regimen components are assumed to be administered by intrathecal injection during a single outpatient appointment.
- Patients are assumed to be eligible to receive HSCT pre-relapse and/or post-relapse. The precise resource use assumptions relating to the HSCT procedure and the initial 2-year follow-up period are not clear from the CS[1] or the source material cited therein.[48, 49] From 2 years after HSCT, patients in post-HSCT follow-up receive 100mg/day cyclosporine indefinitely, but do not incur any further costs associated with visits to health care practitioners. The proportion of patients remaining in HSCT follow-up is assumed to decline over time according to the estimated proportion of patients surviving.
- All patients who relapse receive salvage chemotherapy using FLAG-IDA. This regimen is assumed to be comprised of: filgrastim 0.005mg/Kg (9 days treatment per cycle); fludarabine 30mg/m^2 (5 days treatment per cycle); cytarabine 2,000mg/m^2 (5 days treatment per cycle), and idarubicin 8mg/m^2 (3 days treatment per cycle). The model assumes that 16.8 inpatient days are required to administer this regimen per FLAG-IDA cycle (cycle duration not reported in the CS[1]). Thirty-seven percent of patients who receive one round of salvage chemotherapy are assumed to subsequently receive a further round of the same regimen.

Within the standard care group, the model assumes the following treatment pathway:

- All patients receive chemotherapy whilst relapse-free for up to 2 years unless they undergo pre-relapse HSCT (at which point, chemotherapy is assumed to be discontinued), relapse or die. This treatment is costed according to the maintenance regimen for the non-transplanted population of the UKALL14 trial.[19] This regimen is assumed to be comprised of: (i) vincristine (IV, 1.4mg/m^2 once every 13 weeks); (ii) methotrexate (intrathecal, 12.5mg once every 13 weeks); (iii) prednisolone (oral, 60mg/m^2 5 times every 13 weeks); (iv) mercaptopurine (oral, 75mg/m^2 daily) and (v) methotrexate (oral, 20mg/m^2 weekly).
- HSCT is modelled using the same approach as in the blinatumomab group.
- Salvage chemotherapy is modelled using the same approach as in the blinatumomab group.

The application of different RFS and OS time-to-event curves leads to different trajectories through the main model health states, which when combined with assumptions regarding HSCT use and associated health losses and costs, produce different profiles of total costs and health outcomes for the two treatment groups. Incremental cost-effectiveness is calculated in a pairwise fashion as the difference in costs divided by the difference in QALYs for blinatumomab and standard care.

5.2.3 Key structural assumptions employed within the company's model

The company's model employs the following structural assumptions:

- All patients enter the model in the relapse-free health state.
- HRQoL is principally determined by relapse status, sojourn time in the relapse-free state and treatment group (the latter is driven largely by the treatment-related MRD response rate).
- Blinatumomab is assumed to be continued for up to four six-weekly cycles. Adjunctive prophylaxis against CNS relapse is assumed to be continued for up to nine quarterly cycles, or until HSCT, incidence of relapse, or death.
- Standard care chemotherapy is assumed to be continued for up to eight quarterly cycles, or until HSCT, incidence of relapse, or death.
- The RFS hazard is assumed to follow a Gompertz distribution in both groups (using an approach which is analogous to fitting models independently to each treatment group).
- The OS hazard is assumed to follow a log normal mixture cure model in both groups (which allows a different cure fraction but has the same standard parametric model parameters between the treatment groups).
- The probability of undergoing pre-relapse HSCT is assumed to be constant with respect to time.
- If a patient does not relapse, they are assumed to only be eligible to receive HSCT within the first four years of entering the model.
- Prior to the point at which the proportion of patients who are relapse-free is less than or equal to the cumulative proportion of patients who received a HSCT pre-relapse, all patients who relapse are assumed to have not received a pre-relapse HSCT; after this point, all patients who relapse are assumed to have received a pre-relapse HSCT.

5.2.4 Evidence used to inform the company's model parameters

The main groups of model parameters and the evidence sources used to populate these are summarised in Table 23. These are discussed in further detail in the subsequent sections.

Table 23: Evidence sources used to inform company's model parameters

Parameter type	Parameter	Source(s)
Time-to-event parameters	RFS - blinatumomab	BLAST PAS subgroup[1]
	RFS - standard care	Historical control study DCAS with ATT weights[1]
	OS - blinatumomab	BLAST PAS subgroup[1]
	OS - standard care	Historical control study DCAS with ATT weights[1]
Probability RFS event is death	RFS death probability - blinatumomab	BLAST PAS subgroup[1]
	RFS death probability – standard care	Historical control study DCAS with ATT weights[1]
HRQoL	Health utility – relapse-free ≤5 years	GLM/GEE regression based on BLAST data[1]
	Health utility – relapse-free >5 years (excluding additional HRQoL decrement for cured population)	Kind *et al*[50]
	Health utility – relapsed	Logistic regression using matched patients from BLAST and TOWER subgroups[1]
	QALY loss - HSCT (time-dependent)	Kurosawa *et al*[51]
	QALY loss – proximity to death	GLM/GEE regression based on BLAST data[1]
	Utility decrement for cured population – exposure to radiotherapy, chemotherapy, and HSCT	Assumption based on BLAST GLM/GEE[1] and Kind *et al*[50]
Mean dosing	Proportion of patients receiving blinatumomab dose during each treatment cycle (up to 4 doses)	BLAST[1]*
Probability of receiving HSCT	Probability of receiving pre-relapse HSCT	Calibrated to 4-year data from BLAST (blinatumomab) and historical control (standard care)
	Probability of receiving post-relapse HSCT	Estimated using BLAST and Study NCT02003612 (same probabilities used in each group)
Resource use and costs	Inpatient and outpatient resource use for standard of care and patients discontinuing blinatumomab	Face-to-face interviews with clinical experts (n=2)[1]
	HSCT procedure and subsequent follow-up (0-24 months)	NHS Blood and Transplant.[48] Cyclosporine costs taken from the British National Formulary (BNF)[25]
	Maintenance chemotherapy (standard care group)	Based on subgroup of UKALL14.[19] Unit costs taken from eMIT[52]
	Salvage chemotherapy	NICE TA450[53] (blinatumomab for relapsed/refractory ALL)
	Terminal care costs	King's Fund and Marie Curie reports[54, 55]
	Prophylaxis against CNS relapse for patients receiving blinatumomab	eMIT[52]
	Blinatumomab acquisition cost (including PAS)	Amgen[1]
	Unit costs for visits, appointments, hospitalisations, laboratory tests, radiological tests and AEs	NHS Reference Costs 2015/16[56]

** assumes 50% drug costs for those discontinuing within each cycle*

PAS – primary analysis set; DCAS – direct comparison analysis set; ATT - average treatment effect on the treated; GLM/GEE – generalised linear model/generalised estimating equation; eMIT – Electronic Market Information Tool; TA – technology appraisal

Time-to-event analysis

The company fitted parametric survival curves to time-to-event data from the BLAST PAS and the ATT weighted historical control DCAS. RFS for patients in the blinatumomab group was defined as the interval from the date of first blinatumomab treatment for MT103-203 patients from BLAST until haematological relapse or death (whichever occurred first). In order to avoid an immortal time bias (whereby a patient experiences an event before they are at risk within the study), the RFS interval for the historical comparator patients was adjusted to exclude patients with a time to relapse of less than 14 days (the median time between MRD detection and first blinatumomab dose for BLAST patients); the baseline date for patients within the historical comparator study was set equal to their MRD detection date plus 14 days. OS outcomes for patients in BLAST and the historical control study also relate to these same baseline timepoints, but include only death as an event.

A large range of survival models were fitted to the available RFS and OS data, including: (i) standard parametric models, (ii) restricted cubic spline (RCS) models, and (iii) mixture/non-mixture cure models (see Table 24). For most of the model types considered, the company fitted joint models which include a treatment effect covariate (an HR or constant acceleration factor; referred to in the CS as "restricted" models) and independent models which include treatment group interaction terms for every distributional parameter and are thus equivalent to fitting separate models to the treatment and control groups (referred to in the CS as "unrestricted" models). In addition, the cure models include both unrestricted and restricted model forms as well as third model type which allows a different cure fraction (θ) for the two groups, but the standard model parameters are otherwise the same for the remaining uncured population. This "cure" model form therefore implies that treatment group affects the likelihood of achieving a cure only, whilst for patients who are not cured, the time-to-event distribution is the same for both the standard care and blinatumomab treatment groups. For the RCS models, three variations were considered according to whether splines were fitted to the log cumulative hazard, log cumulative odds, or the inverse normal survival distribution. These are referred to by the company as the RCS Weibull, the RCS log logistic and the RCS log normal, respectively. Although it was not clear from the CS, the code provided by the company following the clarification process[17] (question B4) suggests that all RCS models assume one knot (where an increasing number of knots indicates a more flexible model). Thirty-eight models were fitted to the available RFS data. The same model forms were fitted to the OS data, however three of these (the gamma mixture cure, the gamma mixture cure [unrestricted] and the gamma non-mixture cure [unrestricted]) failed to converge, hence 35 models were fitted to the OS data.

Table 24: Summary of parametric models fitted to RFS and OS data

Standard parametric models	Flexible parametric models	Cure models
Exponential	RCS log logistic (R)	Gamma mixture cure*
Generalised F (R)	RCS log logistic (U)	Gamma mixture cure (R)*
Generalised F (U)	RCS log normal (R)	Gamma mixture cure (U)
Generalised gamma (R)	RCS log normal (U)	Gamma non-mixture cure
Generalised gamma (U)	RCS Weibull (R)	Gamma non-mixture cure (R)
Gompertz (R)	RCS Weibull (U)	Gamma non-mixture cure (U)*
Gompertz (U)	Piecewise exponential	Log normal mixture cure
Log logistic (R)		Log normal mixture cure (R)
Log logistic (U)		Log normal mixture cure (U)
Log normal (R)		Log normal non-mixture cure
Log normal (U)		Log normal non-mixture cure (R)
Weibull (R)		Log normal non-mixture cure (U)
Weibull (U)		Weibull mixture cure
		Weibull mixture cure (R)
		Weibull mixture cure (U)
		Weibull non-mixture cure
		Weibull non-mixture cure (R)
		Weibull non-mixture cure (U)

** Model presented for RFS analysis only*
R – restricted; U – unrestricted; RCS – restricted cubic spline

According to the CS, model discrimination was undertaken based on the consideration of five factors: (i) internal consistency; (ii) goodness-of-fit statistics; (iii) visual fit; (iv) evidence relating to underlying treatment effect, and (v) consistency with external data. The CS does not provide any information regarding the use of clinical judgement to assess the clinical plausibility of the extrapolated portions of the actual fitted parametric curves or their associated hazard functions.

Internal consistency of the RFS and OS models related to two considerations. Firstly, in instances in which the OS model under consideration and the selected base case RFS curves cross (thereby presenting a logically inconsistency), the OS model was excluded from further consideration. Secondly, the CS states that OS models were preferred if the difference in expected post-relapse survival gain between treatment groups was *"relatively small,"*, although little detail is provided in the CS regarding how this judgement was made.

Goodness-of-fit of the RFS and OS models was assessed using the Bayesian Information Criterion (BIC). According to the CS,[1] this measure was selected because it *"penalises overly complex models and its use mitigates risk of overfitting statistical noise in the tails of the observed distributions"* (CS,[1] page 94). Akaike Information Criterion (AIC) statistics for the fitted models were not presented.

Evidence relating to the underlying treatment effect between groups was based on consideration of counterfactual Kaplan-Meier survival plots (whereby estimated treatment effects are applied to the baseline Kaplan-Meier function) and examination of Schoenfeld residuals.[57] Other diagnostic plots (e.g., log cumulative hazard plots or their equivalents) were not presented.

External validity was assessed through comparison of predicted model outcomes with adjusted data from a meta-analysis of studies assessing the association between MRD status and clinical outcomes including EFS and OS in adults with ALL (Berry *et al*[18]). The data from Berry *et al* were used *"to assess the external validity of the RFS and OS distributions used in the model as well as the magnitude of the increase in RFS and OS that would be expected given the effect of blinatumomab on MRD response"* (company's clarification response,[17] question B12).

Given the wide range of parametric models included in the model-fitting process, the company considered only the five best fitting RFS models, determined according to their BIC; all other RFS models were excluded at this point. Similarly, the company considered only the five best fitting OS models which did not produce a logical inconsistency when viewed alongside the selected deterministic base case RFS curve. The other criteria for model choice described above were therefore considered only for these five best-fitting RFS and OS models. The ERG notes a lack of clarity within the CS regarding the company's subjective judgements of "good", "moderate" and "poor" in relation to these other model selection criteria. The company's clarification response[17] (question B16) provides additional detail and describes a "good" fit as *"the two curves are virtually the same, with no systematic over or under estimation"*, and a "poor" fit as *"the two curves are substantially different with apparent systematic over or underestimation over some range of the curve."*

Table 25 presents the BIC statistics for the 38 fitted RFS models. Table 26 presents the BIC statistics for the 35 fitted OS models. The five best fitting (and in the case of OS, logically consistent) models taken forward for further consideration by the company are highlighted in bold in each table.

Table 25: BIC statistics – RFS models

Parametric model	Model class	BIC	Considered further in the CS?
Exponential	Standard	1321.743	No
Generalised F (R)	Standard	1230.616	No
Generalised F (U)	Standard	1244.548	No
Generalised gamma (R)	Standard	1229.286	No
Generalised gamma (U)	Standard	1240.161	No
Gompertz (R)	**Standard**	**1222.061**	**Yes**
Gompertz (U)	**Standard**	**1225.587**	**Yes**
Log logistic (R)	Standard	1228.876	No
Log logistic (U)	Standard	1234.358	No
Log normal (R)	**Standard**	**1227.202**	**Yes**
Log normal (U)	Standard	1232.716	No
Weibull (R)	Standard	1257.919	No
Weibull (U)	Standard	1260.687	No
RCS log logistic (R)	**Flexible parametric**	**1225.662**	**Yes**
RCS log logistic (U)	Flexible parametric	1236.037	No
RCS log normal (R)	Flexible parametric	1229.012	No
RCS log normal (U)	Flexible parametric	1239.741	No
RCS Weibull (R)	Flexible parametric	1230.052	No
RCS Weibull (U)	Flexible parametric	1236.607	No
Piecewise exponential	Flexible parametric	1265.064	No
Gamma mixture cure	Cure	1236.985	No
Gamma mixture cure (R)	Cure	1233.307	No
Gamma mixture cure (U)	Cure	1244.343	No
Gamma non-mixture cure	Cure	1231.214	No
Gamma non-mixture cure (R)	Cure	1233.447	No
Gamma non-mixture cure (U)	Cure	1244.415	No
Log normal mixture cure	Cure	1233.42	No
Log normal mixture cure (R)	Cure	1228.875	No
Log normal mixture cure (U)	Cure	1234.392	No
Log normal non-mixture cure	Cure	1227.46	No
Log normal non-mixture cure (R)	Cure	1229.115	No
Log normal non-mixture cure (U)	Cure	1234.655	No
Weibull mixture cure	Cure	1235.512	No
Weibull mixture cure (R)	Cure	1234.439	No
Weibull mixture cure (U)	Cure	1238.882	No
Weibull non-mixture cure	**Cure**	**1227.299**	**Yes**
Weibull non-mixture cure (R)	Cure	1230.72	No
Weibull non-mixture cure (U)	Cure	1235.785	No

BIC – Bayesian Information Criterion; R – restricted; U – unrestricted; RCS – restricted cubic spline

Table 26: BIC statistics – OS models

Parametric model	Model class	BIC	Considered further in the CS?
Exponential	Standard	1197.457	No
Generalised F (R)	Standard	1176.196	No
Generalised F (U)	Standard	1190.688	No
Generalised gamma (R)	Standard	1173.349	No
Generalised gamma (U)	Standard	1183.772	No
Gompertz (R)	Standard	1181.63	No
Gompertz (U)	Standard	1187.016	No
Log logistic (R)	Standard	1179.883	No
Log logistic (U)	Standard	1185.326	No
Log normal (R)	Standard	1173.671	No
Log normal (U)	Standard	1179.173	No
Weibull (R)	Standard	1197.723	No
Weibull (U)	Standard	1201.822	No
RCS log logistic (R)	Flexible parametric	1169.497	No
RCS log logistic (U)	Flexible parametric	1180.351	No
RCS log normal (R)	Flexible parametric	1171.037	No
RCS log normal (U)	Flexible parametric	1181.938	No
RCS Weibull (R)	Flexible parametric	1169.987	No
RCS Weibull (U)	Flexible parametric	1180.608	No
Piecewise exponential	Flexible parametric	1196.289	No
Gamma mixture cure	Cure	Failed to converge	No
Gamma mixture cure (R)	Cure	Failed to converge	No
Gamma mixture cure (U)	Cure	1194.837	No
Gamma non-mixture cure	Cure	1177.058	No
Gamma non-mixture cure (R)	Cure	1182.231	No
Gamma non-mixture cure (U)	Cure	Failed to converge	No
Log normal mixture cure	**Cure**	**1173.187**	**Yes**
Log normal mixture cure (R)	Cure	1177.834	No
Log normal mixture cure (U)	**Cure**	**1182.969**	**Yes**
Log normal non-mixture cure	Cure	1171.676	No
Log normal non-mixture cure (R)	Cure	1176.964	No
Log normal non-mixture cure (U)	**Cure**	**1182.057**	**Yes**
Weibull mixture cure	**Cure**	**1188.202**	**Yes**
Weibull mixture cure (R)	Cure	1193.661	No
Weibull mixture cure (U)	Cure	1197.174	No
Weibull non-mixture cure	**Cure**	**1183.034**	**Yes**
Weibull non-mixture cure (R)	Cure	1188.552	No
Weibull non-mixture cure (U)	Cure	1192.722	No

BIC – Bayesian Information Criterion; R – restricted; U – unrestricted; RCS – restricted cubic spline

For RFS, the five best fitting models were: (i) Gompertz restricted; (ii) Gompertz unrestricted; (iii) RCS log-logistic; (iv) log normal, and (v) Weibull non-mixture cure. Table 27 summarises the company's judgements regarding model selection for these five best-fitting RFS models. The unrestricted Gompertz was selected for use in the base case analysis *"due to its good statistical fit, visual fit and external validity"* (CS,[1] page 101).

Table 27: Summary of model selection criteria for 5 best-fitting RFS models (adapted from CS Table 37)

Distribution	Δ BIC	Cure fraction	Treatment effect	Visual fit	External validity	Company comments
Gompertz (R)	--	Blin: 48.5% SC: 16.0%*	Moderate	Moderate	Good	Counterfactual plots suggest proportional hazards may overestimate long-term benefit of blinatumomab.
Gompertz (U)	**3.53**	**Blin: 39.5% SC: 17.2%***	**--**	**Good**	**Good**	**Good visual fit, statistical fit, and external validity.**
RCS log logistic (R)	3.6	Blin: 0% SC: 0%	Good	Moderate	Poor	Proportional odds model. Underestimates benefit of blinatumomab relative to external data.
Log normal (R)	5.14	Blin: 0% SC: 0%	Good	Poor	Poor	Accelerated failure time model. Poor visual fit, underestimates benefit of blinatumomab relative to external data.
Weibull non-mixture (cure)	5.24	Blin: 47.8% SC: 15.8%	Moderate	Moderate	Good	Treatment effect† parameterised as a cure model, but also follows proportional hazards. Counterfactual plots suggest proportional hazards may overestimate long-term benefit of blinatumomab.

** Not parameterised as cure models*
† The ERG is unclear about the meaning of treatment effect in this context as the cure model uses the same standard model parameters but a different cure fraction between groups
BIC - Bayesian Information Criterion

For OS, the five best fitting models which do not intersect the deterministic base case unrestricted Gompertz RFS curves were: (i) log normal mixture cure; (ii) log normal non-mixture cure unrestricted; (iii) log normal mixture cure unrestricted; (iv) Weibull non-mixture cure, and (v) Weibull mixture cure. Table 28 summarises the company's judgements regarding model selection for these OS models. The log normal mixture cure was selected for inclusion in the base case analysis as it had a *"much better statistical fit than the other distributions considered"* (CS,[1] page 108).

Table 28: Summary of model selection criteria for 5 best-fitting logically consistent OS models (adapted from CS Table 40)

Distribution	Δ BIC	Cure fraction	Treatment effect*	Visual fit	External validity	Δ PRS (years)	Company comments
Log normal mixture (Cure)	**--**	**Blin: 45.3% SC: 21.3%**	**--**	**Good**	**Moderate**	**-0.70**	**Best-fitting distribution among those consistent with base-case RFS. Large difference in BIC versus next best-fitting distribution.**
Log normal non-mixture (Cure, U)	8.87	Blin: 45.3% SC: 19.3%	--	Good	Moderate	-0.69	Poor statistical fit.
Log normal mixture (Cure, U)	9.78	Blin: 46.6% SC: 21.0%	--	Good	Moderate	-0.65	Poor statistical fit.
Weibull non-mixture (Cure)	9.85	Blin: 42.8% SC: 23.8%	Good	Good	Moderate	-1.58	Poor statistical fit. Treatment effect counterfactual plots are supportive of proportional hazards. Large difference in PRS.
Weibull mixture (Cure)	15.02	Blin: 46.8% SC: 24.9%	--	Good	Moderate	-1.11	Poor statistical fit. Large difference in PRS.

** The ERG is unclear how this could be assessed for cure models and notes that the fields for the log normal mixture cure and Weibull non-mixture cure models are blank*
BIC - Bayesian Information Criterion; PRS – post-relapse survival

Figure 7 presents a comparison of the empirical RFS Kaplan-Meier curves and the unrestricted Gompertz RFS models. Figure 8 presents a comparison of empirical OS Kaplan-Meier curves and log normal mixture cure OS models.

Figure 7: Comparison of empirical RFS Kaplan-Meier curves and RFS unrestricted Gompertz models

Figure 8: Comparison of empirical OS Kaplan-Meier curves and OS log normal mixture cure models

Proportion of RFS events that are deaths

The probability that an RFS event is death is assumed to differ between the treatment groups, based on the BLAST PAS and the ATT-weighted historical control DCAS. As shown in Table 29, 47.1% and 8.5% of RFS events were estimated to be deaths in the blinatumomab and standard care groups, respectively. The CS notes that the higher rate of deaths for blinatumomab may reflect: (i) the more frequent use of HSCT in BLAST; (ii) a *"notable"* proportion of BLAST patients undergoing transplants from mismatched donors thereby leading to greater risks of infection and death, and (iii) potentially incomplete reporting of HSCT receipt in BLAST.

Table 29: Percentage of RFS events which were deaths (reproduced from CS Table 38)

RFS events	BLAST (blinatumomab)		Historical control (standard care)	
	n	%	n	%
Unweighted				
Death	16	47.1%	14	10.7%
Relapse	18	52.9%	117	89.3%
Total	34	100.0%	131	100.0%
ATT-IPTW				

Death	16	47.1%	10.4	8.5%
Relapse	18	52.9%	112	91.6%
Total	34	100.0%	122.3	100.0%
ATE-IPTW				
Death	13.8	40.2%	13	10.1%
Relapse	20.5	59.8%	115.6	90.0%
Total	34.3	100.0%	128.5	100.0%

ATT - average treatment effect on the treated; ATE - average treatment effect; IPTW - inverse probability of treatment weighting

Health-related quality of life

Utility values for the pre-relapse states were based on EQ-5D utility values for patients included in the BLAST PAS (n=[redacted]). The company fitted a generalised linear model/generalised estimating equations (GLM/GEE) regression model with EQ-5D utility as the dependent variable, and covariates for baseline utility, a patient-level indicator variable of MRD response during cycles 1 or 2, a time-dependent indicator variable for on versus off treatment, and a time-dependent indicator variable for death within 6 months.[1] Patients without baseline assessments or any follow-up assessments were excluded from the model. In addition, utility assessments on or after relapse were also excluded from the analysis. A total of 63 patients from the BLAST PAS contributed data to the GLM/GEE model.

The CS states that post-relapse utility assessments in BLAST were limited and were unlikely to be representative of health utility during the entire post-relapse period. Instead, post-relapse utility estimates were based on an ATT matching analysis of the 63 BLAST PAS patients and patients recruited into the TOWER trial of blinatumomab in Ph- R/R BCP-ALL. The CS states that relapsed patients in the CR1 population of BLAST can be considered to be similar to patients in the TOWER trial who did not receive prior salvage therapy and who were not refractory at baseline. A utility value of 0.692 was estimated using this approach. The ERG notes that the precise methods used to generate this value are unclear due to the limited reporting in the CS and the redaction of utility estimates from the Appraisal Committee papers for TA450.[53]

HRQoL decrements associated with HSCT were based on a cross-sectional survey of 524 patients with acute leukaemia (75% acute myeloid leukaemia [AML], 25% ALL) in Japan (Kurosawa *et al*[51]). All patients undergoing HSCT are assumed to experience utility decrements of 0.17, 0.01, and 0.02 during years 1, 2, and 3-5 after HSCT, respectively, based on the differences in the mean utility value at these time points versus >5 years post-HSCT reported by Kurosawa *et al*. The company's model assumes that no further transplant-specific HRQoL decrement is applied 5-years post-HSCT.

A further HRQoL decrement of 0.02 is applied to the general population health utility values to reflect long-term effects of exposure to radiotherapy, chemotherapy, and HSCT. The CS states that this decrement was based on half the difference between the average utility value for blinatumomab patients

in the RFS state, off therapy, and with MRD response (0.842) and the age- and sex-weighted mean population norms for patients between the ages of 35 and 55 (0.877).

Table 30 summarises the health utility values included in the company's model.

Table 30: Health state utilities applied in company's model

Health state	Utility	95% CI	Derivation
Blinatumomab, on-treatment, relapse-free, >6 months prior to death, cycle 1†	0.792	(0.699, 0.886)	Sampling of utility coefficients from the GLM/GEE model, the MRD response rate and baseline utility from the 63 BLAST PAS patients with data
Blinatumomab, on-treatment, relapse-free, >6 months prior to death, cycle 2+†	0.832	(0.789, 0.872)	
Blinatumomab, off-treatment, relapse-free, >6 months prior to death, cycle 1†	0.802	(0.708, 0898)	
Blinatumomab, off-treatment, relapse-free, >6 months prior to death, cycle 2+†	0.842	(0.798, 0.883)	
Standard care, relapse-free, >6 months prior to death	0.806	(0.718, 0.895)	
Blinatumomab and standard care, post-relapse >6 months prior to death	0.692	(0.688, 0.695)	Estimated from logistic regression of TOWER and the 63 BLAST PAS patients with data
General population utility decrement*	-0.02	N/a (constant)	Based on mid-point between utility from BLAST for RFS off-treatment, with MRD response and age- and sex-weighted general population norms[50]
HSCT utility decrement 1-12 months	-0.170	(-0.366, 0.026)	Estimated based on difference in utility from >5 years post-transplant and prior timepoints[51]
HSCT utility decrement 13-24 months	-0.010	(-0.096, 0.076)	
HSCT utility decrement 25-60 months	-0.020	(-0.085, 0.045)	
HSCT utility decrement 61 months+	0.000	N/a - constant	Assumption

** Decrement applied to all age-adjusted utility values*
† CrI generated by the ERG using the company's model
GLM/GEE – generalised linear model/generalised estimating equation; CI – confidence interval

Mean blinatumomab acquisition

Drug acquisition costs for blinatumomab were provided by the company. The company has a Patient Access Scheme in place for blinatumomab resulting in a price of [redacted] for one 38.5µg vial. The model assumes that one vial includes a single dose of useable medication (28µg blinatumomab). The model assumes that patients receive up to four cycles of blinatumomab at a mean dose of 28µg per day for 28 days, followed by 14 days off treatment. This dosing schedule is based on the anticipated marketing authorisation for blinatumomab,[24] rather than the dose used in BLAST (15µg/m^2).[1] Within the BLAST PAS, the mean body surface area (BSA) was 1.89m^2 which leads to a mean dose of 28.4µg,

hence there is no difference in cost between the regimen used in BLAST and the regimen indicated by the marketing authorisation. The model estimates the costs of blinatumomab during each cycle using data on the average of the proportion of patients starting and completing each treatment cycle (see Table 31).

Table 31: Estimated percentage of patients starting and completing each cycle of blinatumomab

Blinatumomab treatment cycle	Patients starting cycle	Patients completing cycle	Assumed treatment proportion in each cycle
1	[redacted]	[redacted]	[redacted]
2	[redacted]	[redacted]	[redacted]
3	[redacted]	[redacted]	[redacted]
4	[redacted]	[redacted]	[redacted]

Blinatumomab administration and associated costs

The model assumes that the administration of blinatumomab is associated with costs relating to inpatient infusions, the pump used to deliver blinatumomab, and outpatient appointments to change the pump bag when treatment is delivered in a home setting.

During the first and second cycles of blinatumomab treatment, patients are assumed to receive 4 days and 2 days of inpatient treatment, respectively; no inpatient days are assumed to be spent delivering blinatumomab during cycles 3 or 4. The cost per day of administering blinatumomab in an inpatient setting was estimated to be £685.56. This value was based on NHS Reference Costs 2015/16[56] and was calculated as the finished consultant episodes (FCE) weighted average of unit costs divided by mean inpatient days for currency codes SA24G-J.

The pump used to deliver blinatumomab was estimated to cost £1,795 and was assumed to have a lifespan of 5 years. The daily cost of the pump was calculated assuming that the pump was used every day during its lifespan. An additional annual maintenance cost of £90 was assumed.

It was assumed that patients require an outpatient visit to change the bag in the pump every 4 days spent receiving blinatumomab in the outpatient setting. These visits were assumed to cost £211.99 per visit, based on NHS Reference Costs 2015/16[56] (outpatient, currency code SB15Z).

Costs associated with prophylaxis against CNS relapse given alongside blinatumomab

In the blinatumomab group, the model assumes that patients receive one outpatient visit per cycle to deliver prophylaxis against CNS relapse (methotrexate, cytarabine and dexamethasone) at a cost of £265.02 (derived from NHS Reference Costs 2015/16,[56] outpatient visit, code SB13Z - Deliver more

Complex Parenteral Chemotherapy at First Attendance). Table 32 summarises the prophylaxis acquisition costs applied in the company's model. For every regimen component, the dose was calculated based on the protocol and the mean BSA in the BLAST PAS. Costs were then calculated assuming that vials and tablets would be perfectly split.

Table 32: Costs associated with prophylaxis against CNS relapse included in the company's model

Treatment	Administration method	Unit size	Tablet/ vial size	Unit cost	Source
Methotrexate	Intrathecal	1000mg	1	£6.63	CMU[52]
Cytarabine	Intrathecal	2000mg	1	£6.60	CMU[52]
Dexamethasone	Intrathecal	3.3mg	10	£2.42	CMU[52]

CMU – Commercial Medicines Unit; mg - milligram

Standard care chemotherapy acquisition

Drug acquisition costs for the standard care group are summarised in Table 33. Standard care chemotherapy was assumed to follow the maintenance regimen for the non-transplanted population of the UKALL14 trial.[19] This regimen is assumed to be discontinued upon receipt of HSCT. Unit costs for all therapies were taken from the Commercial Medicines Unit (CMU) Electronic Marketing Information Tool (eMIT).[52] The model assumes vial sharing with no wastage of pills for oral treatments. For each regimen component, the dose was calculated based on the UKALL14 protocol and the mean BSA in the BLAST PAS. Costs were then calculated assuming that vials and tablets would be perfectly split.

Table 33: Drug acquisition costs applied in the standard care group

Treatment	Administration method	Unit size	Tablet/ vial size	Unit cost	Source
Vincristine	IV	2.0mg	5	£29.26	CMU[52]
Prednisolone	Oral	5mg	28	£0.41	CMU[52]
Mercaptopurine	Oral	50mg	25	£49.15	BNF[25]
Methotrexate	Oral	2.5mg	100	£4.39	CMU[52]
Methotrexate	Intrathecal	1000mg	1	£6.63	CMU[52]

CMU - Commercial Medicines Unit; BNF – British National Formulary; mg - milligram

Standard care chemotherapy administration

In the standard care group, the model assumes that patients receive two outpatient visits per cycle for IV administration of vincristine and intrathecal administration of methotrexate. For intrathecal methotrexate, the cost of administration was assumed to be £265.02, based on NHS Reference Costs 2015/16[56] (outpatient visit, code SB13Z - Deliver more Complex Parenteral Chemotherapy at First Attendance). For vincristine, the cost of administration was assumed to be £304.30, again based on NHS Reference Costs 2015/16[56] (outpatient visit, code SB14Z - Deliver Complex Chemotherapy, including Prolonged Infusional Treatment, at First Attendance). Patients are assumed to self-administer the oral components of the regimen.

Resource use associated with standard care and discontinued blinatumomab

The mean number of additional inpatient and outpatient visits (over and above drug administration visits) for patients who receive standard care and for those who discontinue blinatumomab (according to MRD response status) are based on estimates from face-to-face interviews with two UK experts.[1] These resource use estimates were combined with arm-specific MRD response rates. The MRD response rate in the blinatumomab arm (83.6%) was taken from the subgroup of the BLAST PAS who had an MRD response within the first two cycles of blinatumomab treatment. No data were available on delayed MRD response in the standard care group; the company's model assumes that MRD response for patients receiving standard care is 8.0% based on expert advice that *"this proportion is no greater than 10%"* (CS,[1] page 95). The resource use estimates applied in each treatment group are summarised in Table 34. The costs presented in this table are not applied to patients who are still receiving blinatumomab.

Table 34: Inpatient and outpatient resource use per month by MRD status and associated monthly resource use

Services	Face-to-face interview (n=2)		Resource use applied in each treatment group		Unit cost
	MRD +	MRD-	Discontinued blinatumomab	Standard care	
Inpatient days	1.75	0.06	0.337	1.615	£685.56
Haematologist - outpatient	2.000	1.500	1.918	1.960	£166.03[a]
Radiologist – outpatient	0.417	0.250	0.390	0.404	£51.35[b]
Other specialist – outpatient	0.500	0.250	0.459	0.480	£162.84[c]
General physician - outpatient	0.750	0.417	0.695	0.723	£36[d]

a NHS Reference Costs 2015/16,[56] consultant led face-to-face follow up. Currency code WF01A. Service code 303
b NHS Reference Costs 2015/16,[56] consultant led face to face follow up. Currency code WF01A. Service code 812
c NHS Reference Costs 2015/16,[56] consultant led face to face follow up. Currency code WF01A. Service code 370
d Curtis and Burns 2016
MRD+ - molecular evidence of blasts in the bone above 1 in 10,000, MRD- molecular evidence of blasts in the bone below 1 in 10,000,

Costs associated with HSCT

The model assumes that patients may receive HSCT prior to relapse and/or following relapse. The company's model assumes that patients who are relapse-free may undergo HSCT for up to four years after initiation of treatment with blinatumomab or standard care chemotherapy. The model uses data on the cumulative 4-year probability of having undergone pre-relapse HSCT from the BLAST PAS ([redacted]) and the ATT-weighted historical control DCAS ([redacted]) to inform the blinatumomab and standard care groups, respectively. The modelled 6-monthly probability of receiving HSCT was calibrated such that the predicted cumulative probability of having undergone pre-relapse HSCT at 48 months matches the observed cumulative probabilities.

In the post-relapse population, the model uses four probability inputs as well as the treatment-specific RFS curve to determine the per-cycle probability of receiving post-relapse HSCT. Two probabilities are used to estimate the probability of receiving a post-relapse HSCT: (i) the probability of having a post-relapse HSCT conditional on the patient not having had a pre-relapse HSCT (probability=0.20); (ii) the probability of having a post-relapse HSCT conditional on the patient having previously had a pre-relapse HSCT (probability=0.16). The exact methods and evidence used to estimate these parameters are not clear from the CS.[1] The remaining two probabilities relate to the probability that an RFS event is death in each treatment group (as described in Table 29). These probabilities were estimated from the BLAST PAS for the blinatumomab arm and the historical control DCAS for the standard care arm.

The model predictions for the mean number of HSCTs per patient are summarised in Table 35. As shown in the table, the company's model suggests that the mean number of HSCTs is higher in the

blinatumomab group than the standard care group (mean HSCTs blinatumomab versus standard care -0.79 versus [redacted]).

Table 35: Mean number of HSCTs per patient predicted by the company's model

Treatment group	Mean number of HSCTs per patient		
	Pre-relapse	Post-relapse	Total
Blinatumomab	[redacted]	[redacted]	[redacted]
Standard care	[redacted]	[redacted]	[redacted]

Salvage chemotherapy costs

The salvage chemotherapy regimen is assumed to be FLAG-IDA. The cost of this regimen was estimated to be £16,175 (uplifted to 2015/16 prices), based on the cost estimates reported in NICE TA450.[53] The model assumes that 37% of patients who receive one line of salvage therapy also receive a second line of salvage therapy; this results in a total cost of £21,905 per patient receiving salvage therapy.

Terminal care costs

The model assumes that at the end of life, patients spend 8 weeks (56 days) receiving hospital care. The cost of care (uplifted to 2015/16 prices) was estimated to be £157.74 per day.[55] The mean cost of terminal care was estimated to be £8,834 per patient.

Model evaluation methods

The CS presents the results of the economic analysis in terms of the incremental cost per QALY gained for blinatumomab versus standard care. Results are presented for both the deterministic and probabilistic versions of the model. The CS also includes the results of probabilistic sensitivity analysis (PSA), deterministic sensitivity analyses (DSAs) and scenario analyses. The results of the PSA are presented in the form of a cost-effectiveness plane and cost-effectiveness acceptability curves (CEACs), based on 10,000 Monte Carlo simulations. The results of the DSAs are presented in the form of a tornado diagram for specified model parameters (based on their 95% confidence limits). Alternative scenario analyses are also reported to explore the use of ATE weights, alternative choices of RFS and OS curves, alternative assumptions regarding long-term excess mortality, duration of blinatumomab benefit, and alternative assumptions regarding the probability that an RFS event is death, HSCT use, probability of cure, HRQoL, costs, discount rates and the model time horizon. The distributions applied in the company's PSA are summarised in Table 36. The ERG notes that several uncertain parameters are held fixed at their mean values and some of the choices of distribution and derived standard errors are not appropriate.

Table 36: Distributions applied in company's probabilistic sensitivity analyses

Parameter type	Parameter	Distribution	ERG comment
Patient characteristics	Age, sex, BSA and weight	Fixed	-
Time-to-event parameters	RFS – blinatumomab	Bootstrap	No details provided regarding how the bootstrap procedure was undertaken. It is unclear whether uncertainty in the IPTW weights was included
	RFS – standard of care	Bootstrap	
	OS – blinatumomab	Bootstrap	
	OS – standard of care	Bootstrap	
Probability RFS event is death	RFS death probability - blinatumomab	Beta	-
	RFS death probability – standard care	Beta	-
HRQoL	RFS health utility model baseline	Log normal	Distribution is not bounded by zero and 1.0
	RFS health utility GLM/GEE model parameters (intercept, baseline, off-treatment relapse-free, MRD response and terminal decrement)	Multivariate normal	-
	Health utility - relapsed	Log normal	Distribution is not bounded by zero and 1.0
	QALY loss (time-dependent) – HSCT	Normal	Distribution is not bounded by zero. The HRQoL decrements for HSCT includes positive values in the PSA; this is illogical.
	General population utilities	Fixed	These values are subject to uncertainty
	Utility decrement – exposure to radiotherapy, chemotherapy, and HSCT	Fixed	These values are subject to uncertainty
Mean dosing	Proportion of full blinatumomab dose received (up to 4 doses)	Beta	-
Probability of receiving HSCT	Probability of receiving HSCT pre-relapse	Beta	The per-cycle probability, rather than the 4-yearly probability has been included in the PSA
	Probabilities of receiving HSCT post-relapse	Beta	-
Resource use and costs	Inpatient and outpatient resource use for standard of care and discontinued blinatumomab	Fixed	These values are subject to uncertainty
	HSCT procedure and subsequent follow-up (0-24 months)	Log normal	SE arbitrarily assumed to be 25% of mean
	Maintenance chemotherapy (standard care group)	Fixed	SE arbitrarily assumed to be 25% of mean
	Salvage chemotherapy	Log normal	SE arbitrarily assumed to be 25% of mean
	Terminal care costs	Log normal	SE arbitrarily assumed to be 25% of mean

Parameter type	Parameter	Distribution	ERG comment
	Prophylaxis against CNS relapse for patients receiving blinatumomab	Fixed	-
	Blinatumomab acquisition cost (including PAS)	Fixed	-
	Cost of pump use to deliver blinatumomab	Log normal	SE arbitrarily assumed to be 25% of mean
	Cost of inpatient visits for blinatumomab	Log normal	SE arbitrarily assumed to be 25% of mean
	Unit costs for visits, appointments, hospitalisations, laboratory tests, radiological tests and AEs	Log normal	SE arbitrarily assumed to be 25% of mean. Given that these are based on NHS Reference Costs, SEs could have been calculated using reported interquartile ranges.

SE – standard error

Company's model verification and validation methods

The CS[1] details extensive efforts taken to verify the correct implementation of the model and to ensure the accuracy of the model inputs against the source material from which these were derived. The CS[1] and the clarification response[17] also mention the use of clinical experts to inform certain assumptions within the model (e.g. around the plausibility of cure).

Company's model results

Table 37 presents the central estimates of cost-effectiveness derived from the company's model. Based on the probabilistic version of the model (assuming the unrestricted Gompertz function for RFS and the log normal mixture cure model for OS), blinatumomab is expected to generate an additional 2.85 QALYs at an additional cost of £84,456 compared with standard care: the corresponding incremental cost-effectiveness ratio (ICER) for blinatumomab versus standard care is £29,673 per QALY gained. The deterministic version of the company's model produces a similar ICER of £28,524 per QALY gained for blinatumomab versus standard care.

Table 37: Company's base case cost-effectiveness results – blinatumomab versus standard care (original submitted model)

Probabilistic model					
Option	**QALYs**	**Costs**	**Inc. QALYs**	**Inc. costs**	**Incremental cost per QALY gained**
Blinatumomab	6.96	[redacted]	2.85	£84,456	**£29,673**
Standard care	4.11	[redacted]	-	-	-
Deterministic model					
Option	**QALYs**	**Costs**	**Inc. QALYs**	**Inc. costs**	**Incremental cost per QALY gained**
Blinatumomab	7.10	[redacted]	2.95	£84,259	**£28,524**
Standard care	4.14	[redacted]	-	-	-

Figure 9 and Figure 10 present the results of the company's PSA in the form of a cost-effectiveness plane and CEACs, based on a re-run of the company's original submitted model. Assuming a willingness-to-pay (WTP) threshold (λ) of £20,000 per QALY gained, the company's model suggests that the probability that blinatumomab produces more net benefit than standard care is 0.10. Assuming a WTP threshold of £30,000 per QALY gained, the probability that blinatumomab produces more net benefit than standard care is estimated to be 0.53.

Figure 9: Cost-effectiveness plane – blinatumomab versus standard care (adapted from company's model)

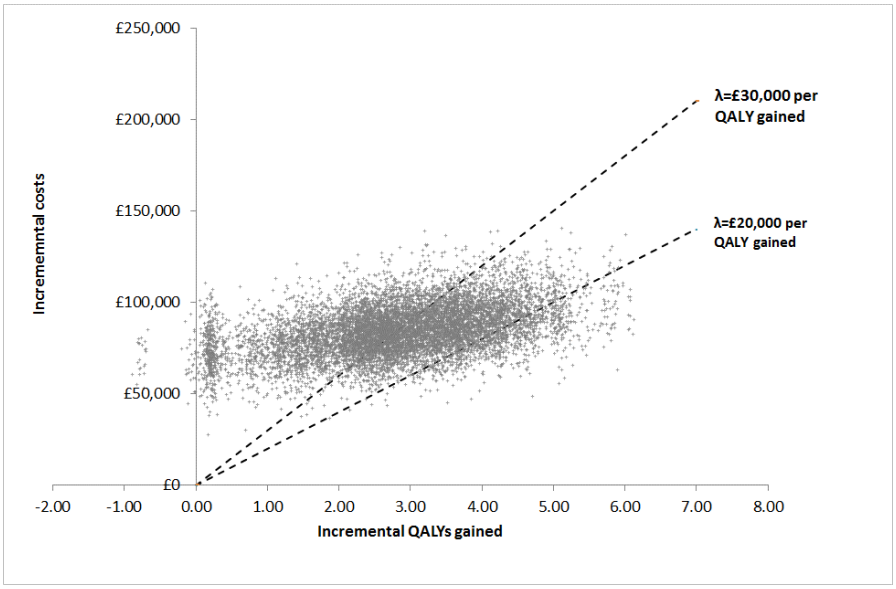


Figure 10: Cost-effectiveness acceptability curves – blinatumomab versus standard care (adapted from company's model)

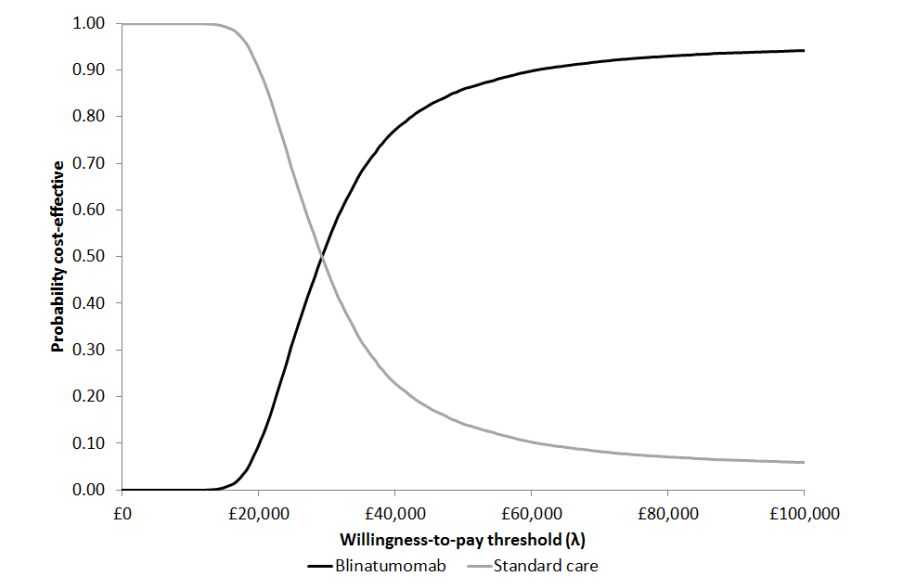


Figure 11 presents the results of the company's DSAs. The DSAs indicate that the five most influential model parameters relate to: (i) the blinatumomab OS cure fraction; (ii) the standard care OS cure fraction; (iii) the proportion of patients in the blinatumomab group who receive HSCT; (iv) the duration of blinatumomab therapy, and (v) the cost of HSCT.

Figure 11: Deterministic sensitivity analysis results – blinatumomab versus standard care (reproduced from company's model)

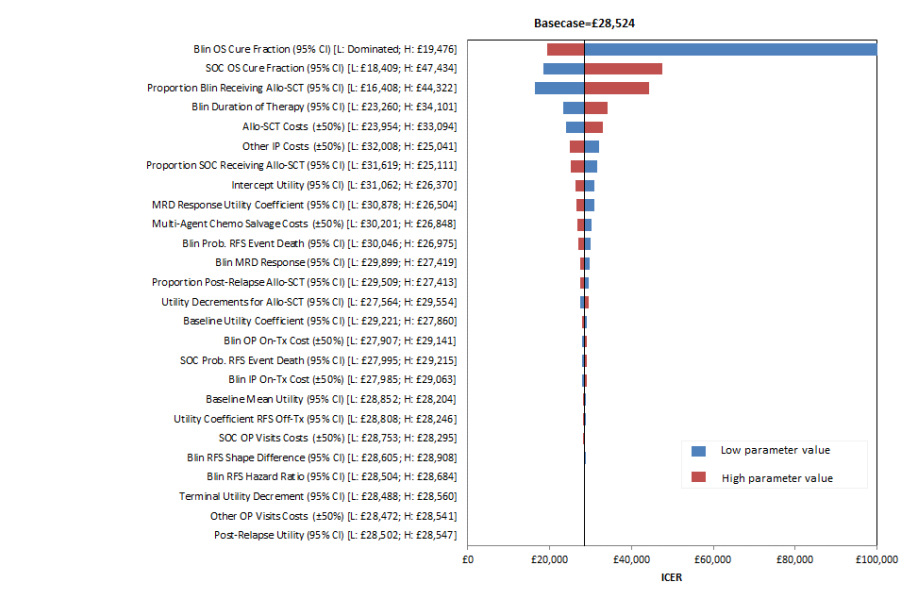


Table 38 presents the results of the company's scenario analyses. The ICER for blinatumomab versus standard care appears to be generally robust to most of the scenarios tested, although the ICER is greater than £30,000 per QALY gained for many scenarios tested. The ICER appears to be particularly influenced by the use of ATE weights, the duration of therapy, the use of the health care resource use survey, inflating OS and RFS outcomes in both groups and the cure fraction.

Table 38: Scenario analysis results – blinatumomab versus standard care (adapted from CS Table 62)

No.	Scenario	Inc. QALYs	Inc. costs	ICER
1	Base case	2.95	£84,259	**£28,524**
1	ATE weights	2.39	£81,370	**£33,999**
2	Alternative extrapolation methods	3.31	£83,064	**£25,081**
3	Unfavourable - RFS RCS log logistic (R), OS RCS Weibull (R)	2.74	£83,874	**£30,647**
4	2-fold increase long-term excess mortality	3.35	£84,300	**£25,199**
5	6-fold increase long-term excess mortality	2.69	£84,234	**£31,274**
6	Duration of benefits = 60 months	2.44	£84,263	**£34,559**
7	Inpatient costs with on-treatment inpatient days from BLAST	2.95	£89,235	**£30,209**
8	Inpatient costs with on-treatment inpatient days from blinatumomab label	2.95	£84,405	**£28,574**
9	23.55% of blinatumomab RFS events are deaths	2.95	£90,548	**£30,698**
10	HRU data from online survey	2.95	£105,376	**£35,673**
11	Cumulative probability of pre-relapse HSCT same for blinatumomab as for standard care	3.00	£49,403	**£16,479**
12	ALL-related costs applied to end of model time horizon	2.95	£80,302	**£27,185**
13	0% MRD response rate for standard care	2.96	£82,537	**£27,892**
14	15% MRD response rate for standard care	2.95	£85,766	**£29,080**
15	No disutility for long-term survivors	3.01	£84,259	**£27,979**
16	0.04 disutility for long-term survivors	2.90	£84,259	**£29,091**
17	Standard care RFS utility = blinatumomab off-treatment RFS utility	2.93	£84,259	**£28,722**
18	ALL-related utilities and costs only to 36 months	2.92	£87,100	**£29,866**
19	ALL-related utilities and costs only to 48 months	2.94	£85,364	**£29,056**
20	Model timeframe = 30 y	2.85	£84,126	**£29,552**
21	Model timeframe = 60 y	2.95	£84,259	**£28,524**
22	Annual discount rate for costs and QALYs=1.5%	3.76	£85,119	**£22,639**
23	Limitations relating to generalisability of standard care arm to current practice (RFS and OS survival distribution based on the ATT analysis of the historical cohort study is adjusted upwards by a factor of 15%)	2.22	£80,202	**£36,163**
24	Blinatumomab OS cure fraction = midpoint OS cure fractions (incremental cure fraction halved)	1.61	£78,918	**£49,101**

ATE – average treatment effect; ATT – average treatment effect on the treated; R – restricted; RCS – restricted cubic spline; HRU – health care resource use

Updated model results

In response to minor issues raised by the ERG during the clarification process, the company provided an updated model which included the following amendments: (i) maximum annual mortality risk capped at 100%; (ii) pump costs included for all days after the first inpatient stay; (iii) general population utilities based on Ara and Brazier,[58] and (iv) post-relapse allogeneic HSCT not initiated after 5 years. The updated model results are similar to the company's original base case (see Table 39); the probabilistic ICER for blinatumomab versus standard care is estimated to be £28,655 per QALY gained.

Table 39: Company's base case cost-effectiveness results – blinatumomab versus standard care (updated model submitted following clarification)

Probabilistic model					
Option	**QALYs**	**Costs**	**Inc. QALYs**	**Inc. Costs**	**Incremental cost per QALY gained**
Blinatumomab	7.11	[redacted]	2.92	£83,634	**£28,655**
Standard care	4.19	[redacted]	-	-	-
Deterministic model					
Option	**QALYs**	**Costs**	**Inc. QALYs**	**Inc. Costs**	**Incremental cost per QALY gained**
Blinatumomab	7.23	[redacted]	3.02	£83,800	**£27,779**
Standard care	4.21	[redacted]	-	-	-

5.3 Critical appraisal of the company's health economic analysis

This section presents a critical appraisal of the health economic analysis presented within the CS.[1] Section 5.3.1 details the methods used by the ERG to interrogate and critically appraise the company's submitted health economic analysis. Section 5.3.2 discusses the extent to which the company's analysis adheres to the NICE Reference Case.[26] Section 5.3.3 summarises the ERG's verification of the company's implemented model and highlights inconsistencies between the model, the CS, and the sources used to inform the model parameter values. Section 5.3.4 presents a detailed critique of the main issues and concerns underlying the company's analysis.

5.3.1 Methods for reviewing the company's economic evaluation and health economic model

The ERG adopted a number of approaches to explore, interrogate and critically appraise the company's submitted economic evaluation and the underlying health economic model upon which this was based. These included:

- Consideration of key items contained within published economic evaluation and health economic modelling checklists[59, 60] to critically appraise the company's model and analysis.
- Scrutiny of the company's model by health economic modellers and discussion of issues identified amongst the members of the ERG.
- Double-programming of the deterministic version of the company's model to fully assess the logic of the model structure, to draw out any unwritten assumptions and to identify any apparent errors in the implementation of the model.
- Examination of the correspondence between the description of the model reported within the CS[1] and the company's executable model.
- Replication of the base case results and PSA presented within the CS.[1]
- Where possible, checking of parameter values used in the company's model against their original data sources.

- The use of expert clinical input to judge the credibility of the company's economic evaluation and the assumptions underpinning the model.

5.3.2 Adherence of the company's model to the NICE Reference Case

The company's economic evaluation is generally in line with the NICE Reference Case[26] (see Table 40). The ERG notes that the model excludes relevant patient subgroups which are included in the proposed marketing authorisation and that inevitably there is considerable uncertainty surrounding the results of the analysis due to the observational nature of the data. These issues are discussed in further detail in Section 5.3.4.

Table 40: Adherence of the company's model to the NICE Reference Case

Element	Reference case	ERG comments
Defining the decision problem	The scope developed by NICE	The ERG notes that the model reflects a population of patients who are able to receive chemotherapy; however, blinatumomab represents a potential treatment option for patients who are unable to undergo HSCT or to tolerate chemotherapy. In addition, two further potentially overlapping subgroups of the BLAST study were excluded from the indirect comparison and health economic model: (i) patients with Ph- MRD+ BCP-ALL in CR2+; (ii) patients with Ph+ MRD+ BCP ALL.
Comparator(s)	As listed in the scope developed by NICE	The company's model compares blinatumomab against standard care chemotherapy. The final NICE scope[22] included a second comparator which was defined as "monitor for relapse." This has not been included as an option in the company's model; the ERG notes that this comparator would be relevant to patients who are unable to receive HSCT or to tolerate chemotherapy.
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Health gains accrued by patients are modelled in terms of QALYs gained.
Perspective on costs	NHS and PSS	Whilst not explicitly stated in the CS,[1] the company's economic analysis adopts an NHS and PSS perspective.
Type of economic evaluation	Cost-utility analysis with fully incremental analysis	The company's economic evaluation takes the form of a cost-utility analysis. The results of the analysis are presented in terms of the incremental cost per QALY gained for blinatumomab versus standard care.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	The company's model adopts a 50-year time horizon. By this timepoint, more than 99.9% of the modelled population have died.
Synthesis of evidence on health effects	Based on systematic review	Health outcomes are modelled using IPTW weighted data from the BLAST PAS and the historical control DCAS (both studies are currently unpublished).

Element	Reference case	ERG comments
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults.	HRQoL estimates for the relapse-free state were derived from GLM/GEE regression analyses of patient-reported EQ-5D data collected in the BLAST study.[1] The HRQoL estimate for the post-relapse state was derived from a logistic regression analysis using the TOWER trial and the 63 patients in the BLAST study with HRQoL data.[1] Additional HRQoL estimates are based on the literature[50, 51] and assumptions.
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	
Source of preference data for valuation of changes in HRQoL	Representative sample of the UK population	
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	No additional equity weighting is applied to estimated QALY gains. The CS argues that blinatumomab meets NICE's criteria for a life-extending end of life treatment. The CS also argues that blinatumomab meets many of the criteria for appraisal under the NICE HST framework and should be evaluated taking into account a wider range of criteria about the benefits and costs.
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	Resource components included in the company's model reflect those relevant to the NHS and PSS. Unit costs were valued at 2015/16 prices.
Discount rate	The same annual rate for both costs and health effects (currently 3.5%)	Costs and health effects are discounted at a rate of 3.5% per annum.

PSS – Personal Social Services; HRQoL – health-related quality of life

5.3.3 Model verification and correspondence between the model, the CS and parameter sources

Model verification

The ERG rebuilt the deterministic version of the company's base case model in order to verify its implementation; the results of the model rebuild are shown in Table 41. As shown in the table, the ERG's rebuilt model produces very similar estimates of health gains, costs and cost-effectiveness compared with the company's model. During the process of rebuilding the company's base case economic model, seven minor implementation/programming errors were identified:

(i) The annual general population mortality rate is applied for 1-year intervals defined according to time since model entry, rather than according to patient age. However, the initial patient age is not an integer (initial age = 45.38 years), hence applying the modelled mortality =LOOKUP() function for a full year is incorrect.

(ii) The risk of all-cause death exceeds 1.0 for males patients aged 95 years and older and female patients aged 97 years and older. This error was rectified within the company's updated model provided as part of the company's clarification response[17] (question B42).

(iii) The formula used to calculate the receipt of HSCT at 2 years is subject to minor programming errors.

(iv) The formula used to apply discounting to the cost of other inpatient visits post-relapse in the blinatumomab arm is subject to a programming error caused by the formula being incorrectly offset.

(v) Post-relapse HSCTs were assumed to occur after the 5-year time point; however, elsewhere in the model, ALL-related costs were not applied after 5 years as it was assumed that patients would not relapse beyond this timepoint. This error was rectified within the company's updated model provided as part of the company's clarification response[17] (question B33).

(vi) The application of the utility decrement due to death for each model cycle was calculated based on the number of deaths occurring 6 months into the future. Within the first 5 years of the time horizon, the model assumes that all deaths are ALL-related and should therefore be subject to the utility decrement (based on the GLM/GEE model). The model multiplies the utility decrement calculated from the GLM/GEE by the number of people who were expected to die either within either: (i) the next 27 model cycles, or (ii) before the model time horizon reaches 5 years. This approach is inappropriate, as the utility decrement for a patient who dies within a model time cycle should depend on the patient's current survival probability and their history, rather than events occurring in the future.

(vii) Discounting is incorrectly applied to the HSCT costs due to the use of approximate =LOOKUP() functions used to calculate the discount rate for receipt of HSCT.

The ERG notes that these errors have only a minor impact on the ICER for blinatumomab versus standard care.

Table 41: Comparison of company's original submitted base case model and ERG's rebuilt model including PAS

Option	Company's model			ERG's rebuilt model		
	QALYs	Costs	ICER	QALYs	Costs	ICER
Blinatumomab	7.10	[redacted]	**£28,524**	7.10	[redacted]	**£28,529**
Standard care	4.14	[redacted]	-	4.14	[redacted]	-

QALY – quality-adjusted life year; ICER - incremental cost-effectiveness ratio

Correspondence between the written submission and the model

The implemented model appears to be generally in line with its description within the CS.[1] However, the ERG considers that the logic and implementation of the HSCT sub-models are not well described in the CS. In addition, limited detail is provided regarding the logistic regression of the TOWER and BLAST data used to generate the post-relapse utility value. As individual patient-level data (IPD) were not provided by the company, it was not possible for the ERG to fully verify the implementation of the survival models described in the CS.

Correspondence of the model inputs and the original sources of parameter values

The ERG was unable to locate the company's estimated cost of death within the King's Fund and Marie Curie reports;[54, 55] however, the value used within the model should not have a material impact on the model results. In addition, the ERG could not identify the cost of salvage therapy (£16,175) or the source of the assumption that 37% of patients receiving salvage therapy would receive a subsequent further line of salvage therapy within the Appraisal Committee papers from TA450.[53] As the company produced these analyses for an earlier appraisal, this lack of correspondence is unlikely to be an important issue. Further, the ERG was unable to source the parameter value relating to the proportion of patients who survive 24 months after receiving HSCT (20%). All other parameter values correspond with their original sources.

5.3.4 Main issues identified within the critical appraisal

Box 1 summarises the main issues identified within the ERG's critical appraisal of the company's economic analysis. These issues are discussed in further detail in the subsequent sections.

Box 1: Summary of main issues identified within the company's health economic model

(1) Exclusion of relevant patient groups from the economic analysis
(2) "Monitor for relapse" comparator not included in the model
(3) Use of a model structure which is inappropriate for tracking HSCT
(4) Absence of RCT evidence for blinatumomab versus standard care
(5) Concerns regarding company's approach to RFS/OS model selection
(6) Concerns regarding the robustness of the company's alternative base case (blinatumomab used on relapse for the standard care group)
(7) Questionable reliability of the company's HRQoL estimates
(8) Uncertainty surrounding the proportion of RFS events that are deaths
(9) Unrealistic treatment pathway
(10) Limited sensitivity analysis around alternative parametric functions

(1) Exclusion of relevant patient groups from the economic analysis

The population considered within the company's economic analysis relates to patients with Ph- MRD+ BCP-ALL with first complete haematological remission (CR1). This modelled population is narrower than the anticipated marketing authorisation for blinatumomab,[24] as it excludes three relevant subgroups of patients: (i) patients who are unable to receive HSCT or to tolerate chemotherapy; (ii) patients with Ph- MRD+ BCP- ALL with CR2+, and (iii) patients with Ph+ MRD+ BCP-ALL. The CS[1] argues that blinatumomab should be considered for use in alignment with its full anticipated marketing authorisation (for the treatment of adults with MRD+ BCP-ALL).

In response to a request for clarification (see clarification response,[17] question A2), the company noted that there is limited evidence relating to patients with CR2+. Based on the results of the BLAST study, patients with CR1 and MRD response had better outcomes than patients with CR2 and MRD response, however, those in CR2 and MRD response still gained benefit from blinatumomab (see Table 42). However, the historical control study included only patients with CR1, hence there are no data available for comparison. The CS[1] and the company's clarification response[17] also noted that clinical advice received by the company suggested that blinatumomab would be used as early in the pathway as possible and that *"subsequent use of blinatumomab to treat MRD positivity in later remission states or as a salvage therapy is not anticipated if blinatumomab is used in the aforementioned [first-line] setting."* On this basis, the company argues that the CR1 population is the most appropriate ICER for decision-making. Clinical advisors to the ERG agreed that blinatumomab would be used as early as possible in the treatment pathway. However, the ERG notes that the exclusion of patients with CR2+ reduces the available sample size from the BLAST study (41 of 116 [35.3%] patients had second or third CR).

Table 42: Summary of OS and RFS for blinatumomab-treated patients in CR2 in BLAST (adapted from clarification response question A2)

CR2 subpopulation, BLAST	MRD responders	MRD non-responders
RFS, median (months)	[redacted]	[redacted]
OS, median (months)	[redacted]	[redacted]

The company's clarification response[17] notes that the Ph+ population was not represented in the model as the number of Ph+ patients recruited into BLAST was very small (n=5), and the historical control study did not include these patients. Clinical advisors to the ERG stated that the treatment pathway for Ph+ ALL is markedly different from that for Ph- ALL, as several effective treatment options (specifically, TKIs) are available for these patients.

The CS makes the argument that MRD+ patients who have a high risk of relapse would not solely be monitored for relapse without any active treatment. However, clinical advisors to the ERG suggested

that some patients may not be sufficiently fit to receive HSCT or chemotherapy, but may be able to tolerate blinatumomab. The company's model does not assess the cost-effectiveness of blinatumomab within this population.

The ERG considers that on the basis of the evidence submitted to NICE, it is not possible to make any reliable estimate of the cost-effectiveness of blinatumomab in these excluded population groups.

(2) "Monitor for relapse" comparator not included in the model

The company's model compares blinatumomab against standard care chemotherapy. The final NICE scope[22] listed an additional comparator which was defined as "monitor for relapse"; this option is not considered as a comparator in the company's model (see critical appraisal point 1). The ERG considers that the company's economic analysis should have explored an assessment of the cost-effectiveness of blinatumomab versus monitoring within the subgroup of patients who are unable to receive HSCT or retreatment with chemotherapy, but for whom blinatumomab is an option.

(3) Issues relating to the modelling of HSCT

The model attempts to incorporate the impact of HSCT through two mechanisms: (i) the principal benefits of HSCT in reducing and/or avoiding the risk of relapse and death are implicitly reflected in the RFS and OS time-to-event analyses, and (ii) the QALY losses and costs associated with the HSCT procedure and post-HSCT survival are reflected within two HSCT sub-models. The approach adopted by the company to capture the impact of HSCT is subject to several limitations: (a) the absence of a causal link between HSCT uptake and its impact on RFS and OS outcomes; (b) the model cannot estimate the probability that a patient receives HSCT; (c) the adoption of questionable assumptions regarding HSCT receipt, and (d) the likely underestimation of post-HSCT costs.

(i) Absence of a causal link between HSCT uptake and its impact on RFS and OS outcomes

The model does not include a causal link between the extent of HSCT use and the principal RFS/OS benefits resulting from the use of this intervention. For example, setting the 6-monthly probability of receiving HSCT to zero reduces the HSCT-related costs and QALY losses to zero, however, the RFS and OS outcomes remain unchanged. The absence of a direct structural link between the extent of HSCT use and the benefits and costs accrued as a consequence of HSCT makes it difficult to judge the reliability of this aspect of the model.

(ii) The model cannot estimate the probability that a patient receives HSCT

Given that HSCT is not explicitly incorporated into the company's model structure, it is not possible to track the proportion of patients who undergo HSCT post-relapse (including those patients who undergo more than one transplant). As such, it is not possible to calculate the proportion of people who would

receive an HSCT (although it is possible to estimate the overall number of HSCTs received per patient). As a consequence, this aspect of the model is not transparent and it is difficult to determine whether the assumed use of HSCT is clinically plausible.

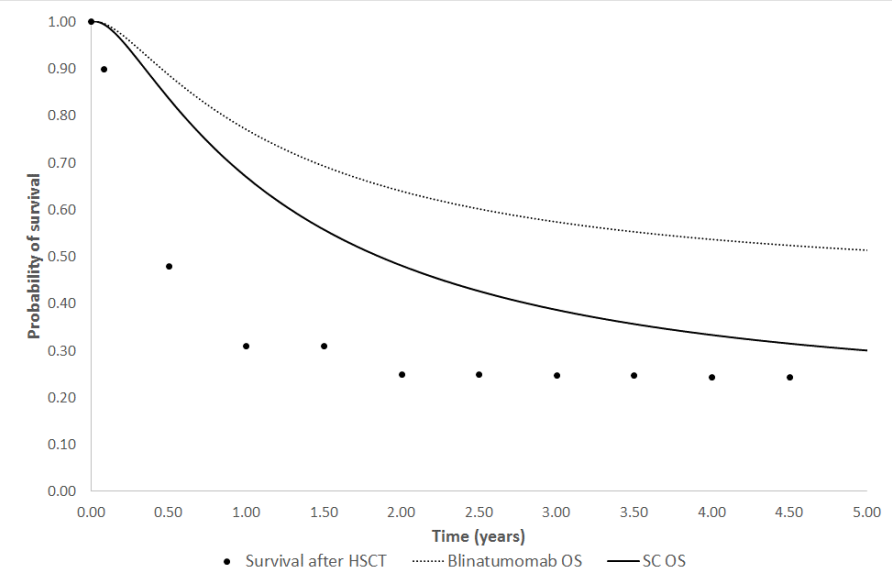
(iii) Adoption of questionable assumptions regarding HSCT receipt

The HSCT post-relapse sub-model makes the following assumptions: (i) HSCTs occur in 6-monthly batches (thereby affecting the cumulative proportion of patients still receiving chemotherapy at each timepoint), and (ii) patients with pre-relapse HSCT do not relapse until all patients without pre-relapse HSCT have relapsed. The first assumption could have been avoided by using the same cycle duration within the HSCT sub-models and the main partitioned survival model. The second assumption could have been avoided only through the use of a different overall model structure which would allow for the tracking of HSCT history across the patient cohorts. Whilst the data used to inform the frequency of post-relapse HSCT is weaker than that for pre-relapse HSCT, the overall frequency of post-relapse HSCT is low in both treatment groups, hence it does not have a substantial impact on the ICER.

(iv) Likely underestimation of post-HSCT costs

The probability of remaining alive post-HSCT over time is approximated using data on survival post-HSCT from NHS Blood and Transplant[48] (it is implicitly assumed that all surviving patients remain in post-HSCT follow-up) and using uplifted general population mortality estimates (from 2-years post-transplant onwards). These data are used to estimate the costs and health losses associated with HSCT and post-transplant care, but do not affect survival gains (see point (3i) above). Figure 12 presents a comparison of the parametric (log normal mixture cure) OS curves for each treatment group and the assumed survival post-HSCT applied in the HSCT sub-models. It should be noted that the modelled OS curves reflect what happens to all patients, including those who receive HSCT as well as those who do not, whilst the NHS Blood and Transplant data reflect survival in an exclusively transplanted cohort. Clinical advisors to the ERG noted that they would expect that, other things being equal, OS would be higher in transplanted patients compared to non-transplanted patients. However, as shown in Figure 12, OS in the transplanted cohort is markedly worse than that for both the blinatumomab and standard care groups. The consequence is that the model appears to include significant benefits in terms of OS due to cure following HSCT, but underestimates both the long-term costs and QALY losses associated with this treatment. Model testing undertaken by the ERG indicates that increasing the costs of post-HSCT follow-up and cyclosporine and increasing the HRQoL decrements associated with HSCT both lead to increases in the ICER for blinatumomab versus standard care.

Figure 12: Comparison of the assumed survival post-HSCT to the company's base case OS curves in both the blinatumomab and standard care arms for someone who received their HSCT within the first cycle of the HSCT sub-model



The ERG notes that in order to explicitly capture the extent of HSCT use pre- and post-relapse, and the costs and benefits accruing as a consequence of those procedures, a different model structure would be required (e.g. a semi-Markov model or a discrete event simulation [DES]). This would allow for tracking of patient histories, however, it would also require a re-analysis of the available time-to-event data to account for competing risks of relapse and death within transplanted and non-transplanted subgroups. The ERG believes that following such an approach would lead to two key benefits: (i) the incorporation of structural links between the use of HSCT and its associated costs and health impacts; (ii) the incorporation of more explicit assumptions regarding the benefits of HSCT (e.g. survival in transplanted and non-transplanted patients) which would improve model transparency and credibility. However, the ERG notes that the available data to populate specific transitions would be limited by very small sample sizes, may be subject to selection bias, and would be associated with considerable uncertainty.

(4) Absence of RCT evidence for blinatumomab versus standard care

As described in Section 4.4, propensity score methods based on IPTW were used to provide adjusted Kaplan-Meier survival curves for the standard care chemotherapy group. Although this is appropriate given the absence of RCT evidence, this introduces an important limitation for all subsequent analyses.

The propensity score weights (and hence the adjusted Kaplan-Meier survival curves) are estimates with associated uncertainty. It is unclear (although unlikely) that this uncertainty has been accounted for in the subsequent model fitting.

(5) Concerns regarding company's approach to RFS/OS model selection

The ERG has concerns regarding the company's approach to model selection. As detailed in Section 5.2, the company fitted a large number of parametric models to the available RFS and OS data. The company then selected the five best fitting RFS curves judged according to their BIC statistics and the five best fitting OS curves judged according to their BIC statistics and whether the given OS model was logically consistent with the final selected RFS function. Other aspects of model choice were considered only for the five best-fitting functions. The ERG notes that the CS does not provide any information regarding the use of clinical judgement to inform decisions regarding the plausibility of the selected RFS and OS functions, however, the company's clarification response[17] (question B7) states that UK clinicians were asked to comment on: (i) the expected survival of patients currently observed in clinical practice (at landmark timepoints); (ii) the appropriateness of assuming a cure at a specific timepoint; (iii) the proportion of patients that may realise a cure given current treatments, and (iv) the magnitude of benefit likely to be derived from obtaining an MRD-negative status. On the basis of the information provided in the CS, it does not appear that clinicians were asked to judge which specific parametric models appear most plausible.

The ERG considers that many of the curves fitted by the company are unnecessary and/or inappropriate. Clinical advice received by the company (see CS[1] page 120 and clarification response[17] question B7) and the ERG suggests that patients who have not relapsed within 5-years may generally be considered to be cured. Therefore, it seems reasonable to assume that a cure model is appropriate from the outset. However, the company's use of BIC to select out a subset of models for further consideration results in a situation whereby two of the "best" RFS models do not predict a cure fraction and are thus clearly inappropriate. It may be the case that other models which fit the data less well during the observed period may produce more plausible extrapolations, however these are excluded from further consideration due to company's application of an initial model selection criterion based on BIC.

The company fitted restricted, unrestricted, cure, cure (restricted) and cure (unrestricted) models to the available time-to-event data. The ERG considers that it would be appropriate to include only unrestricted models from the outset for two reasons. Firstly, whilst it is possible to explore the assumption of proportional hazards/constant accelerated failure over the observed period of the studies included in the analyses, this assumption may not hold within the extrapolated period, hence, the ERG would prefer to exclude models which apply such restrictive assumptions. Secondly, the data for the comparator are weighted but not directly observed and are subject to uncertainty. This uncertainty

appears to have been ignored in the analysis presented by the company, with equal consideration given to the observed BLAST data and the weighted historical control data, with the latter having a larger sample size and so having a higher influence in the resulting model fit statistics. It would therefore be more appropriate to conduct the model fitting separately in both groups.

The ERG considers that these choices around the use of a cure model and the use of treatment effect covariates should have been made *a priori*. The ERG also notes that the company's model selection should have explored the clinical plausibility of the fitted models (based on full models which include the mortality hazard for cured patients).

In addition, the ERG notes that there appears to be some inconsistency between the clinical advice received by the company regarding the likelihood of achieving cure and the time at which the model predicts that such cure occurs. Figure 13 and Figure 14 present the modelled RFS and OS functions for the blinatumomab and standard care groups, respectively. The crosses marked on each RFS/OS curve show the "cure point" predicted by the model, that is, the timepoint at which the hazard for RFS or OS drops below the company's assumed hazard of other-cause mortality within this population. Beyond this timepoint, the model assumes that the only remaining event is other-cause death (uplifted from general population life tables). As shown in both Figure 13 and Figure 14, the timepoint at which the modelled RFS/OS event hazard reverts to that for the uplifted general population mortality differs between the RFS and OS endpoints within the same population. For the blinatumomab group, the model indicates a cure point at 7.28 years for RFS and 8.01 years for OS. Within the standard care group, the difference between the RFS and OS cure points is more pronounced, with cure being modelled from 5.63 years for RFS and 11.00 years for OS. The ERG believes that whilst it is possible that a proportion of patients might achieve cure following relapse (due to downstream HSCT), which may justify the use of models which imply different timepoints for cure (as suggested within the company's clarification response,[17] question B6), it is not clinically plausible to apply models which feature such a large gap between those achieving cure pre- and post-relapse.

Whilst the ERG recognises the difficulties of generating robust survival models given the evidence available, the ERG would have preferred the adoption of a model structure which aligns directly with the clinical input received by the company and the ERG - that patients who have not relapsed within 5-years are considered to be cured. However, the ERG notes that owing to the use of a partitioned survival model, applying the assumption of cure at 5-years to the OS curves produces a bias, as patients who are alive and relapsed at this timepoint gain additional survival benefit. The use of an alternative model structure (e.g. a state transition model) would rectify this problem, but may introduce alternative issues due to small sample sizes and an increased risk of selection bias.

Figure 13: RFS and OS cure points – blinatumomab group

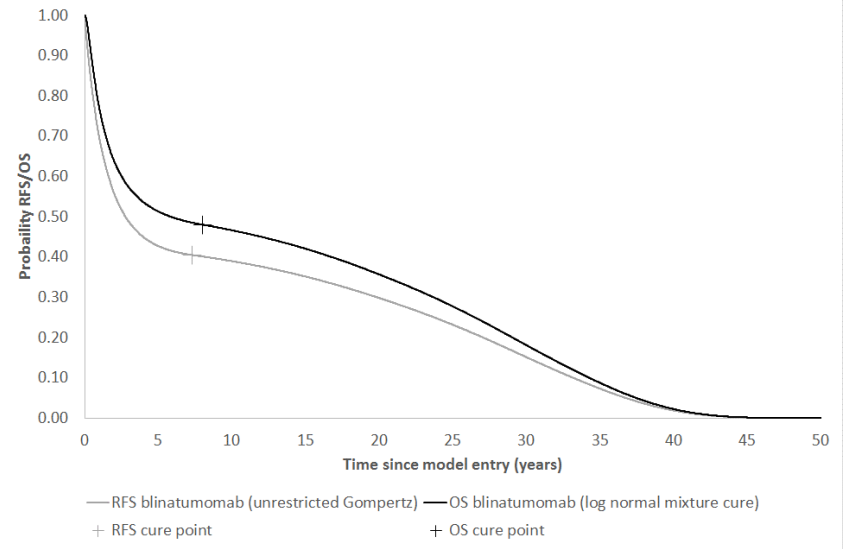
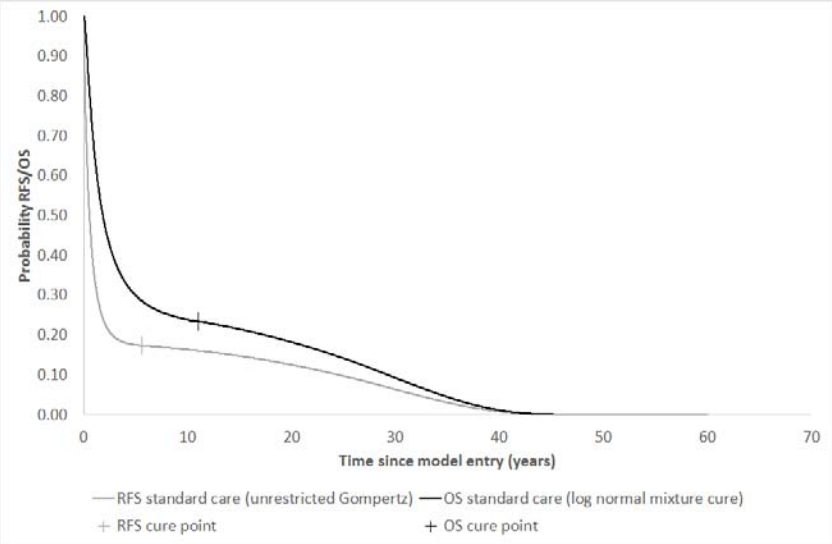


Figure 14: RFS and OS cure points – standard care group



On the basis of clinical advice received by the company and the ERG, the ERG considers that it would be more appropriate to apply a fixed cure point at 5-years.

(6) Concerns regarding the robustness of the company's alternative base case (blinatumomab used on relapse for the standard care group)

The CS[1] presents a "key scenario analysis" (which is referred to as an "alternative base case" in the company's clarification response,[17] question A2) in which blinatumomab is assumed to be used as first salvage therapy for 70% of patients who relapse on standard care chemotherapy. This analysis is based on the incremental survival gains, QALY gains and costs of blinatumomab (versus FLAG-IDA) from TA450;[53] these are added in to the base case total health gains and costs. This analysis produces an ICER for blinatumomab versus standard care of £17,420 per QALY gained (see Table 43).

Table 43: Company's alternative base case (blinatumomab used on relapse for standard care group)

Option	QALYs	Costs	Inc. QALYs	Inc. Costs	Incremental cost per QALY gained
Blinatumomab	7.10	[redacted]	1.91	£33,473	**£17,420**
Standard care	5.19	[redacted]	-	-	-

The ERG considers that this analysis is problematic for two reasons. Firstly, the MRD+ relapsed population within the current model reflects only a subgroup of the relapsed/refractory population within the TA450 model.[53] Secondly, the additional costs and health outcomes assumed to be related to blinatumomab salvage therapy are not structurally related to the OS gains estimated from the company's survival modelling of data from the BLAST PAS and the historical control study DCAS. As such, the ERG considers the results of this analysis to be highly uncertain.

(7) Questionable reliability of the company's HRQoL estimates

The ERG has several concerns regarding the plausibility of the HRQoL estimates assumed within the model.

(i) Relapse-free utility

The GLM/GEE-derived utility values for the RFS state (utility=0.79-0.84 depending on cycle, treatment and whether the patient has discontinued treatment) are similar to the general population utilities reported by Kind *et al*[50] (utility = 0.844). The clinical advisors to the ERG considered the relapse-free utility to be a reasonable reflection of the HRQoL for this population.

(ii) Post-relapse utility

The ERG has some concerns regarding the post-relapse utility value estimated using the logistic regression of the TOWER and BLAST studies.[1, 61] Clinical advisors to the ERG considered that the post-relapse utility estimate of 0.692 appears to be unrealistically high. The CS provides only limited details regarding the derivation of this estimate. During clarification (see clarification response,[17]

question B20), the ERG requested further information regarding the observed EQ-5D estimates in the relapsed population of BLAST. In response, the company stated there were only 8 post-relapse utility assessments, of which 6 assessments were conducted on the day of relapse, 1 assessment was conducted 22 days after relapse, and 1 assessment was conducted 30 days after relapse. The mean utility value for these 8 post-relapse assessments was 0.819 (0.276). The ERG notes that this value is higher than some of the relapse-free utility estimates derived from the company's GLM/GEE model and that this estimate is therefore not reliable. The ERG's clinical advisors suggested that HRQoL in relapsed patients would likely be much lower (estimated utility=0.25 to 0.60), irrespective of whether the patient was fit enough for transplant.

(iii) General population HRQoL

The company's original base case model used the study reported by Kind *et al*[50] to estimate age- and sex-specific general population health utilities. The ERG considers the regression study of Health Survey for England (HSE) data reported by Ara and Brazier[58] to represent a more appropriate source for these parameters, as it includes a larger sample size (Ara and Brazier n=26,729; Kind *et al* n=3,395) and it is more up-to-date (Ara and Brazier, 2010 [based on HSE 2003 and 2006]; Kind *et al* 1999 [based on data collected in 1993]). In response to a request for clarification[17] (question B21), the company updated their model to use HRQoL estimates from Ara and Brazier.[58] The company's clarification response[17] notes that the utility values from Ara and Brazier[58] are generally slightly higher than those based on Kind *et al;*[50] as such, the use of these newer estimates yields a slightly more favourable ICER for blinatumomab versus standard care compared with the company's original base case (ICER using Ara and Brazier=£27,938 per QALY gained; ICER using Kind *et al*=£28,524).

(iv) Decrement associated with exposure to radiotherapy, chemotherapy, and HSCT

The ERG considers that the HRQoL decrement associated with radiotherapy, chemotherapy, and HSCT, which is based on a mid-point value, is essentially arbitrary. The ERG also notes that during the first 5-years post-HSCT, the proportion of this decrement which is attributable to HSCT should already be captured through the QALY losses estimated through the HSCT sub-models.

(8) Uncertainty surrounding the proportion of RFS events that are deaths

The ERG notes that there is a considerable difference in the proportion of RFS death events between the data from BLAST and the ATT-weighted data from the historical control study DCAS (BLAST PAS RFS death probability = 47.1%, historical control DCAS RFS death probability = 8.5%). The CS makes the case that the high probability observed in BLAST may be a consequence of incomplete capture of relapses after transplant in BLAST and mismatched donors resulting in infections.

The ERG agrees that there may be issues surrounding incomplete data collection in BLAST, as the level of censoring is considerably higher than in the historical control study DCAS (BLAST PAS total n=34 events; ATT-weighted historical control study total n=122.3 events). However, it is not clear that the proportion of death events would necessarily decrease with additional follow-up. Furthermore, it is unclear from the CS whether infections caused by mismatched donors and intensive immunosuppression were the cause of death in these patients. The ERG notes that decreasing the RFS death proportion in the blinatumomab group leads to a less favourable ICER.

(9) Unrealistic treatment pathway

The company's model captures a single treatment pathway for the standard care comparator. This is assumed to be comprised of chemotherapy according to the UKALL14 maintenance therapy regimen[19] (vincristine, methotrexate [intrathecal], prednisolone, mercaptopurine and methotrexate [oral]) followed either by HSCT(s) and/or salvage chemotherapy (FLAG-IDA). Clinical advice received by the ERG suggests that the treatment pathways for patients with Ph- MRD+ BCP-ALL are more complex and depend on the patient's level of MRD positivity, patient fitness, their eligibility for allograft (including the availability of matched donors), as well as variability between centres and paediatric and adult haematologists.

Clinical advisors to the ERG provided the following description of the treatment pathway for patients with MRD+ BCP-ALL.

At present, adults aged 16-60 years being treated for ALL with curative intent in the UK will receive intensive chemotherapy that can broadly be described in 4 phases – induction, intensification and consolidation followed by maintenance. Different terms are used in the paediatric protocol although the chemotherapies used are similar. Although there is no routine allografting in the paediatric protocol UKALL2011, the current adult protocol UKALL14 stipulates that most adults receive an allogeneic transplant rather than continue with chemotherapy alone. Allografting usually occurs post intensification in place of consolidation and maintenance.

Patients that have persistent MRD following induction chemotherapy are at an increased risk of relapse and will usually require an allogeneic transplant to have any chance of cure. The exception to this is the younger teenage patients with low levels of MRD $<10^{-3}$ where it may be acceptable to continue with chemotherapy only in some circumstances. The success of allografting in adults is directly linked to the levels of MRD prior to the transplant. Adult patients with persistent MRD $<10^{-3}$ may be cured by an allograft (although this chance is increased if MRD can be reduced to $<10^{-4}$). Those under 40 years of age would be suitable to go straight to a myeloablative transplant at this stage. Those over 40 years of

age would have intensification with high-dose methotrexate as an inpatient before receiving a reduced intensity allogeneic transplant.

Transplantation is unlikely to be curative when the levels of MRD post induction are 10^{-3} or higher. In this situation, patients will require more intensive blocks of salvage chemotherapy as an inpatient in order to try and reduce the levels of MRD prior to allografting. However, these patients may have chemo-refractory disease and may not be able to achieve deeper levels of MRD in which case an early relapse is likely.

Those patients that have persistent MRD and are not able to proceed to an allograft for any reason e.g. no suitable donor, failure to reduce MRD to an acceptable level or poor general fitness, will be given standard chemotherapy in an attempt to prolong life although this strategy is unlikely to be curative.

The ERG's clinical advisors also noted that patients would receive FLAG-IDA as salvage chemotherapy; after failing this regimen, a different regimen would be used.

The ERG therefore has concerns that the company's model does not fully reflect the complexity of current treatment pathways followed by patients in England. Specifically, the ERG notes the following:

- The company's model only includes a single standard care chemotherapy regimen
- The model does not reflect any interplay between patient characteristics (e.g. age, fitness, eligibility for HSCT) and treatments received.
- The company's assumption that patients who fail FLAG-IDA salvage would receive further therapy using this regimen is inappropriate.

(10) Limited sensitivity analysis around alternative parametric functions

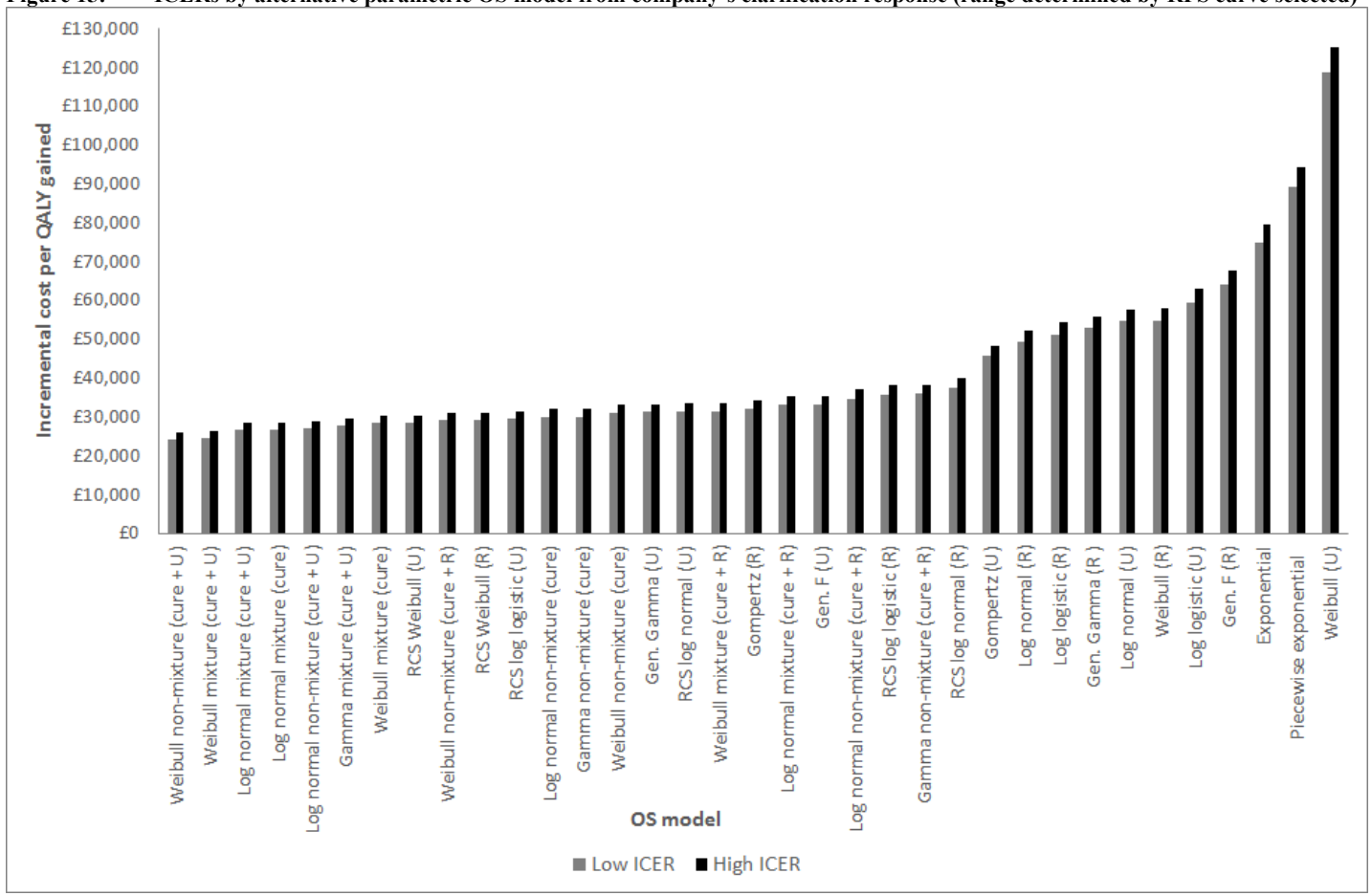
The company fitted 38 separate models to RFS and 35 separate models to OS. Whilst this indicates that there are many possible combinations of potentially plausible RFS and OS models, the CS includes only two additional scenario analyses which explore the impact of using alternative parametric functions for RFS and OS:

(i) RFS and OS distributions changed to restricted Gompertz and unrestricted Weibull non-mixture cure, respectively. The ICER for this scenario is reported to be £25,081 per QALY gained.
(ii) RFS and OS distributions changed to restricted RCS log-logistic and restricted RCS Weibull, respectively. The ICER for this scenario is reported to be £30,647 per QALY gained.

In response to a request for clarification[17] (question B8), the company presented analyses which combine different RFS and OS models across 1,330 different combinations. Figure 15 presents the

distribution of the resulting ICERs according to the selected OS function, with low and high ICERs indicating the impact of assuming different RFS functions given the selected OS function. The Weibull non-mixture cure (unrestricted) OS model produces the lowest ICER (£24,171 per QALY gained); the Weibull (unrestricted) OS model produces the highest ICER of (£125,153 per QALY gained). The highest ICER arising from any OS cure model is £38,076 per QALY gained. The ERG notes that the company's base case ICER is towards the lower end of the range of possible ICERs.

Figure 15: ICERs by alternative parametric OS model from company's clarification response (range determined by RFS curve selected)



5.4 ERG exploratory analyses

ERG exploratory analyses - methods

The ERG undertook eight sets of exploratory analysis. All analyses were undertaken using the deterministic version of the updated model submitted by the company following clarification.[17] As the bootstrap RFS and OS samples for the company's base case model selections were hardcoded into the model, it was not possible to re-run the probabilistic model using alternative RFS/OS functions. Technical details relating to the implementation of these analyses can be found in Appendix 1.

The ERG's analyses include two key exploratory analyses which when combined represent the ERG-preferred model. These analyses are detailed below:

Exploratory analysis 1: Correction of errors. Within this analysis, seven programming errors identified during the ERG's double-programming and model verification exercise were rectified (see Section 5.3.3).

Exploratory analysis 2: Inclusion of a fixed cure point of 5-years. Within this analysis, the hazard of death for all patients surviving beyond 5-years was switched from the hazard predicted by the parametric model to that of the uplifted general population (where a given model has a cure point which manifests at less than 5 years from model entry, this assumption has no effect). This amendment was implemented using existing functionality contained within the company's model. It should be noted that this assumption better reflects clinical judgement and means that the cure point is applied structurally at a fixed timepoint, rather than being determined by the statistical model. However, due to limitations in the company's model structure, the ERG does not consider this analysis to be ideal, as the cure is applied to both RFS and OS functions at the cure point; patients who are alive and have relapsed at 5-years (9% of the blinatumomab group and 13% of the standard care group) will therefore be considered cured, which is not realistic. Given the model structure, it was not possible to relax this assumption. Therefore, this analysis will produce a bias in favour of the standard care group, although the ERG considers the magnitude of this is likely to be small.

Exploratory analysis 3: ERG-preferred model. This analysis combines exploratory analyses 1 and 2. Notwithstanding the uncertainty surrounding the selection of parametric RFS and OS functions, this analysis represents the ERG's preferred model.

In addition, five further sets of exploratory analyses were undertaken using the ERG's preferred model:

Exploratory analysis 4: Exploration of impact of alternative standard care chemotherapy costs. Within this analysis, the drug acquisition costs for standard care chemotherapy were doubled in order to assess

the impact of assuming alternative treatment regimens on the ICER for blinatumomab versus standard care.

Exploratory analysis 5: Exploration of the impact of alternative post-HSCT survival probabilities. Within this analysis, post-HSCT survival was estimated using data on the 100-day mortality rate after allogenic HSCT from BLAST[39] and the uplifted age- and sex-weighted general population mortality rates thereafter. For the first 6-monthly cycle post-HSCT, the probability of death was calculated by adding the 100-day mortality rate from BLAST to the probability of death in the remaining 82.6 days of the cycle using the uplifted general population mortality rates. For all subsequent 6-monthly cycles, the probability of death was estimated using the uplifted general population mortality rates.

Exploratory analysis 6: Exploration of alternative cure fractions for the standard care group. This analysis was undertaken to assess the sensitivity of the model results to the assumed cure fraction for the standard care group. Analyses were undertaken for cure fractions of 25%, 30% and 35%.

Exploratory analysis 7: Exploration of alternative post-relapse HRQoL estimates. Within this analysis, three alternative HRQoL estimates were applied to the post-relapse state in order to explore their impact on the ICER for blinatumomab versus standard care: (i) the observed EQ-5D value for the small number of patients with post-relapse utility assessments in BLAST;[1] (ii) an assumed value of 0.50 and (iii) an assumed value of 0.25.

Exploratory analysis 8: Exploration of the impact of alternative parametric RFS and OS models. Within this analysis, the model was run assuming alternative unrestricted parametric OS and RFS models across a total of 1,330 model combinations. Clinical advisors to the ERG were asked to select their preferred unrestricted OS function and to give reasons supporting their selections (see Figure 16 and Figure 17 for survival plots; full model selection questionnaire presented in Appendix 2).

Figure 16: Predicted cumulative survival probabilities by OS model type (including 5-year cure assumption and mortality risk in cured population) - blinatumomab

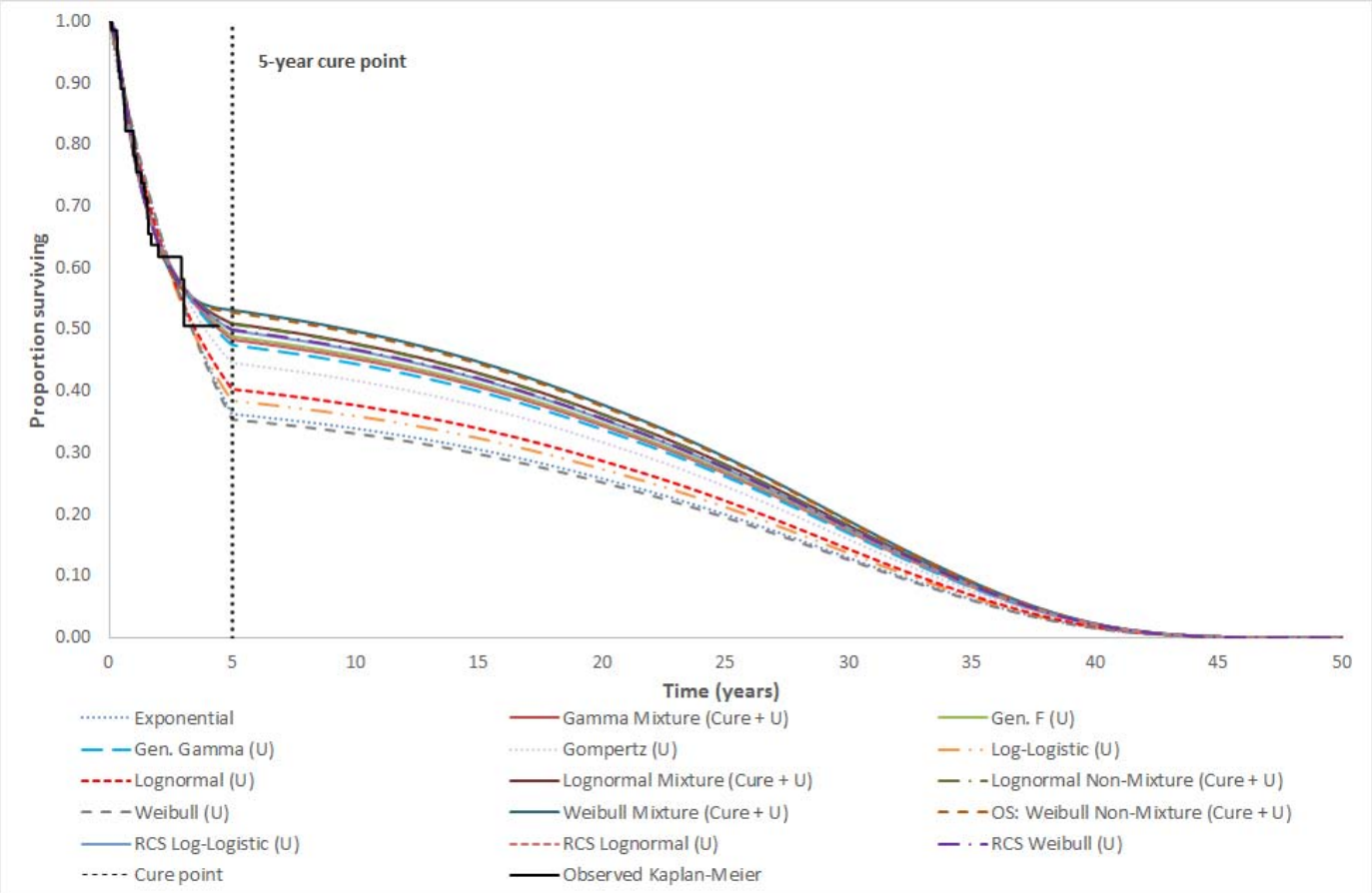
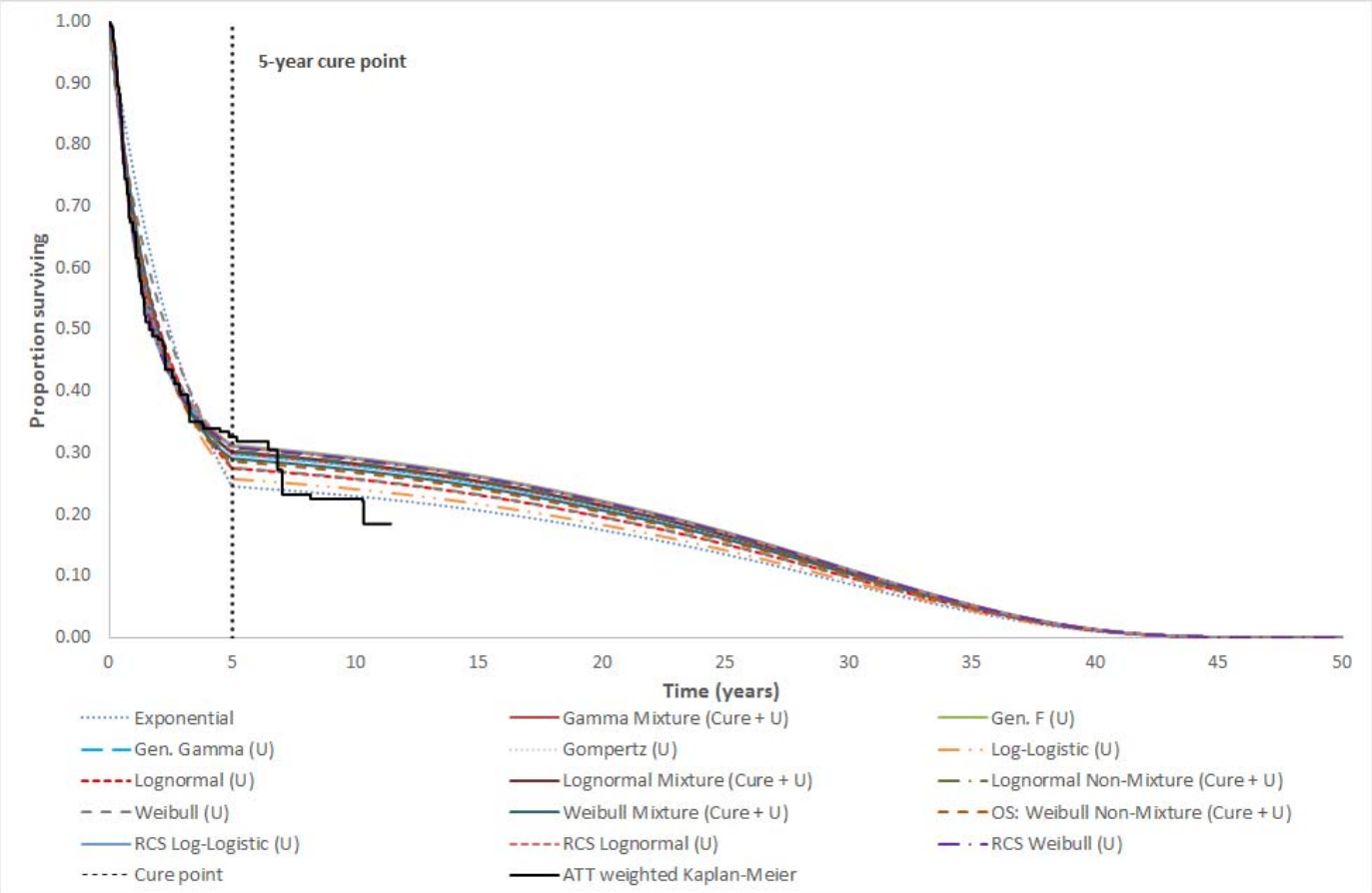


Figure 17: Predicted cumulative survival probabilities by OS model type (including 5-year cure assumption and mortality risk in cured population) – standard care



ERG exploratory analyses - results

ERG exploratory analyses 1-3 – Correction of errors and inclusion of a 5-year cure point

Table 44 presents the results of exploratory analyses 1-3. Analyses 1 and 2 are applied individually to the company's updated model submitted post-clarification; analysis 3 combines both analyses to reflect the ERG's preferred model. As shown in Table 44, the correction of errors has only a minor impact upon the ICER for blinatumomab versus standard care (ICER=£27,717 per QALY gained). The incorporation of an assumption of cure at 5-years also leads to a slightly less favourable ICER for blinatumomab versus standard care (ICER=£30,304 per QALY gained). When these analyses are combined, the deterministic ICER for blinatumomab versus standard care is estimated to be £30,227 per QALY gained.

Table 44: Results of ERG exploratory analyses 1-3 (error correction and inclusion of a 5-year cure point)

Option	QALYs	Costs	Inc. QALYs	Inc. Costs	Incremental cost per QALY gained
Company's base case (updated model)					
Blinatumomab	7.23	[redacted]	3.02	£83,800	**£27,779**
Standard care	4.21	[redacted]	-	-	-
ERG exploratory analysis 1 – Correction of errors identified during model verification					
Blinatumomab	7.21	[redacted]	3.00	£83,264	**£27,717**
Standard care	4.21	[redacted]	-	-	-
ERG exploratory analysis 2 – Cure applied to all surviving patients at 5 years					
Blinatumomab	7.37	[redacted]	2.77	£83,803	**£30,304**
Standard care	4.61	[redacted]	-	-	-
ERG exploratory analysis 3 –Analyses 1 and 2 combined (ERG-preferred model)					
Blinatumomab	7.35	[redacted]	2.75	£83,268	**£30,227**
Standard care	4.59	[redacted]	-	-	-

Further sensitivity analyses undertaken using the ERG-preferred model

Table 45, Table 46, Table 47, Table 48, Table 48 and Figure 18 present additional sensitivity analyses around the ERG's preferred model in order to explore the impact of alternative assumptions of the ICER for blinatumomab versus standard care.

ERG exploratory analysis 4 – Standard care chemotherapy costs doubled

Table 45 presents the results of an analysis in which the costs of standard care chemotherapy were doubled. This analysis suggests that the costs of standard care chemotherapy do not materially impact upon the ICER for blinatumomab.

Table 45: ERG exploratory analysis 4 – Standard care chemotherapy costs doubled (based on the ERG-preferred model)

Option	QALYs	Costs	Inc. QALYs	Inc. Costs	Incremental cost per QALY gained
Blinatumomab	7.35	[redacted]	2.75	£82,222	**£29,848**
Standard care	4.59		-	-	-

ERG exploratory analysis 5 – Use of alternative HSCT survival probabilities

Table 46 presents the results of an analysis in which the probability of remaining alive and in follow-up following HSCT were increased, based on the 100-day mortality rate for blinatumomab and uplifted general population mortality rates. This analysis indicates that the HSCT survival probabilities lead to an increase in the ICER for blinatumomab versus standard care, although the ERG notes there is uncertainty surrounding the survival trajectory of patients undergoing HSCT.

Table 46: ERG exploratory analysis 5 – Use of alternative HSCT survival probabilities (based on the ERG-preferred model)

Option	QALYs	Costs	Inc. QALYs	Inc. Costs	Incremental cost per QALY gained
Blinatumomab	7.29	[redacted]	2.73	£89,302	**£32,667**
Standard care	4.55		-	-	-

ERG exploratory analysis 6 – Use of alternative cure fractions for standard care chemotherapy

Table 47 presents the results of the analyses whereby the cure fraction for the standard care group was set equal to 0.25, 0.30 and 0.35, respectively. The results of the analysis highlight that the cure fraction is a key driver of cost-effectiveness for blinatumomab versus standard care.

Table 47: ERG exploratory analysis 6 – Use of alternative cure fractions for standard care chemotherapy (based on the ERG-preferred model)

Standard care cure fraction	Blinatumomab versus standard care		
	Inc. QALYs	Inc. costs	ICER
Cure fraction = 0.21 (company's base care)	2.75	£83,268	**£30,227**
Cure fraction = 0.25	2.36	£81,402	**£34,465**
Cure fraction = 0.30	1.83	£78,883	**£43,072**
Cure fraction = 0.35	1.30	£76,363	**£58,697**

Exploratory analysis 7 – Impact of alternative post-relapse utility values

Table 48 presents the results of the analyses in which alternative post-relapse utility values are applied. As shown in the table, the post-relapse utility value has a fairly minor impact on the ICER, with lower values resulting in more favourable ICERs for blinatumomab versus standard care. The ERG notes that even at extreme values of post-relapse utility (for example, utility=0.25), the ICER is reduced only by around £3,000.

Table 48: Exploratory analysis 7 – Impact of alternative post-relapse utility values (based on the ERG-preferred model)

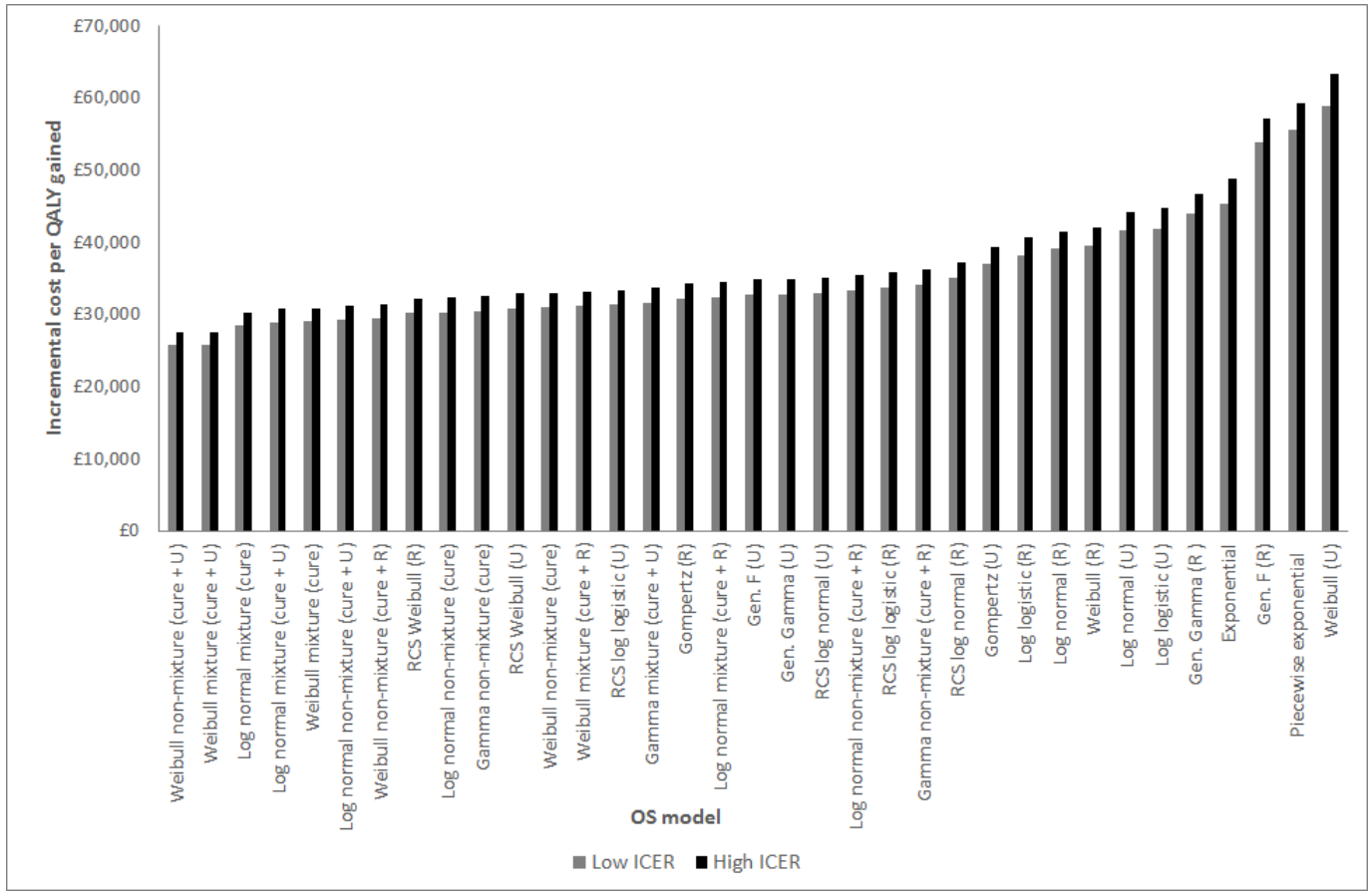
Post-relapse utility value	Blinatumomab versus standard care		
	Inc. QALYs	Inc. costs	ICER
Utility=0.69 (company's base case)	2.75	£83,268	**£30,227**
Utility=0.819 (BLAST post-relapse utility[1])	2.67	£83,268	**£31,157**
Utility=0.50	2.88	£83,268	**£28,930**
Utility=0.25	3.04	£83,268	**£27,395**

ERG exploratory analysis 8 - Impact of alternative parametric RFS and OS models on the ICER for blinatumomab

Figure 18 presents the results of additional analyses of the ERG's preferred model in which a large range of alternative parametric models are assumed for RFS and OS. For each OS model, the range of low and high ICERs reflects the impact of assuming alternative RFS functions. Based on the ERG's preferred model, this exploratory analysis indicates the following:

- The inclusion of the 5-year cure assumption reduces the variation in ICERs across the OS models considered. The ICER for blinatumomab versus standard care ranges from £25,783 per QALY gained (Weibull non-mixture cure model, unrestricted) to £63,265 per QALY gained (Weibull, unrestricted).
- As with the company's analyses presented in Section 5.3, for a given OS model, the RFS function does not generally produce a large range in terms of the highest and lowest ICER. The ICER range for RFS given the selected OS model is typically around £2,000.
- In general, the cure models produce lower ICERs than the other OS functional forms (standard parametric models and RCS models).
- Only the Weibull non-mixture cure model (unrestricted) and the Weibull mixture cure model (unrestricted) produce results in which the full range of ICERs are below £30,000 per QALY gained.
- The ICERs at the lower end of the range for the log normal mixture (cure), log normal mixture (cure, unrestricted), Weibull mixture (cure) and log normal non-mixture (cure, unrestricted) and Weibull non-mixture (cure, unrestricted) are below £30,000 per QALY gained.

Figure 18: ERG exploratory analysis 8 - Impact of alternative parametric RFS and OS models on the ICER for blinatumomab (low-high ICER range determined by RFS curve given the selected OS model)



The clinical advisors to the ERG considered the assumption of a cure point at 5 years to be acceptable and noted that this is in line with data observed in the UKALLXII trial,[62] whereby the Kaplan-Meier OS curves begin to approach an approximate plateau from around year 3.

Advice on the range of plausible statistical models was also consistent. For blinatumomab, the preferred distribution given by one advisor was the generalised gamma. This was selected on the basis that the estimated OS of 50% at 5 years was considered to concord with the observed data from BLAST and study MT103-202. The ERG's second clinical advisor selected the RCS Weibull model and the log normal mixture/non-mixture cure models as their preferred choice based on these providing clinically expected changes in OS between years four and five. It was noted that after four years, the rate of events would be expected to decrease, in line with a cure point at around five years. Models suggesting a steep drop in OS during this interval were considered implausible as these provide predictions with an unrealistic change in the hazard rate at 5 years when combined with the elevated general population mortality estimates. This led both clinical advisors to dismiss the four lowest predicting models (log normal, log logistic, exponential and Weibull). The first advisor also stated that models which provided a more favourable OS profile than the RCS Weibull (Weibull mixture cure [unrestricted], Weibull non-mixture cure [unrestricted], log normal mixture cure [unrestricted], log normal non-mixture cure [unrestricted]) were unlikely to be plausible.

For standard care chemotherapy, the Weibull mixture cure was selected as the preferred distribution by the first clinical advisor, based on its predicted 5-year OS probability. Models between the log normal and RSC Weibull were considered to be plausible. The second clinical advisor selected the RSC Weibull as the preferred distribution based on the fit to the (ATT-weighted) observed data up to five years. Both clinical advisors expressed uncertainty in the clinical plausibility of the observed drop in OS from year 6 onwards, which did not reflect their experience in clinical practice by which time very few events would be expected.

Table 49: ERG clinical advisors' list of potentially plausible OS models (preferred model highlighted in bold)

OS model	Clinical advisor 1	Clinical advisor 2
Blinatumomab	RCS Weibull (U) RCS log logistic (U) Generalised F (U) RCS log normal (U) Gamma mixture cure (U) **Generalised Gamma (U)** Gompertz (U)	Log normal mixture cure (U) Log normal non-mixture cure (U) **RCS Weibull (U)** RCS log logistic (U) Generalised F (U) RCS log normal (U) Gamma mixture cure (U) Generalised gamma (U) Gompertz (U)
Standard care	RCS Weibull (U) Log normal non-mixture cure (U) Log normal mixture cure (U) Generalised gamma (U) Gamma mixture cure (U) Gompertz (U) **Weibull mixture cure (U)** OS: Weibull non-mixture cure (U) Weibull (U) Log normal (U)	**RCS Weibull (U)** No clear range given due to similarity of curves

U – unrestricted; RCS – restricted cubic spline

Table 50 summarises the ICER ranges associated with the three OS models preferred by the ERG's clinical advisors. The clinical advisors' three preferred OS models (Generalised gamma [unrestricted], RCS Weibull [unrestricted] and Weibull mixture cure [unrestricted]) result in ICERs in the range £25,810 per QALY gained to £34,904 per QALY gained.

Table 50: ICERs associated with clinical advisors' preferred OS functions (low-high ICER range determined by RFS curve given the selected OS model)

OS model	Low ICER	High ICER
Generalised gamma (U)	£32,800	£34,904
RCS Weibull (U)	£30,868	£32,857
Weibull Mixture (Cure + U)	£25,810	£27,492

U – unrestricted

5.5 Discussion

The CS includes a systematic review of published economic evaluations of treatments for adult ALL patients with MRD-positivity after treatment together with a *de novo* health economic analysis of blinatumomab versus standard care chemotherapy in patients with Ph- MRD+ BCP-ALL. The company's review did not identify any published economic evaluations of blinatumomab in this indication.

The company's *de novo* partitioned survival model assesses the cost-effectiveness of blinatumomab versus chemotherapy (based on the UKALL14 maintenance regimen) in patients with Ph- MRD+ BCP-

ALL in CR1. Incremental health gains, costs and cost-effectiveness of blinatumomab are evaluated over a 50-year time horizon from the perspective of the NHS and PSS. The company's model is comprised of a main structure which reflects RFS and OS outcomes, as well as two linked sub-models which are intended to estimate additional costs and HRQoL decrements associated with HSCT given before and/or after relapse. The main model structure includes three health states: (1) relapse-free; (2) post-relapsed and (3) dead. The model parameters were informed by analyses of time-to-event data (RFS and OS) from the company's IPTW weighted analysis of the BLAST PAS and the ATT-weights historical control study DCAS. RFS is modelled using an unrestricted Gompertz distribution (using an approach which is analogous to fitting models independently to each treatment group), whilst OS is modelled using a log normal mixture cure model (whereby the treatment effect is applied only to the cure fraction parameter). HRQoL is assumed to be principally determined by relapse status, time spent in the relapse-free state and treatment received; utility estimates were derived from a GLM/GEE model fitted to EQ-5D data collected in BLAST, a propensity matching analysis of the BLAST and TOWER blinatumomab studies, as well as other literature and assumptions. Resource use estimates and costs were based on data collected in BLAST, the UK ALL14 treatment protocol, routine cost sources, clinical opinion and other literature.

Based on the probabilistic version of the model (assuming the unrestricted Gompertz function for RFS and the log normal mixture cure model for OS), blinatumomab is expected to generate an additional 2.85 QALYs at an additional cost of £84,456 compared with standard care: the corresponding ICER for blinatumomab versus standard care is £29,673 per QALY gained. The deterministic version of the company's model produces a similar ICER of £28,524 per QALY gained for blinatumomab versus standard care. Assuming a WTP threshold (λ) of £20,000 per QALY gained, the company's model suggests that the probability that blinatumomab produces more net benefit than standard care is 0.10. Assuming a WTP threshold of £30,000 per QALY gained, the probability that blinatumomab produces more net benefit than standard care is estimated to be 0.53. Following the clarification process, the company submitted a revised model which addressed some of the minor concerns initially raised by the ERG; this updated model generated an ICER for blinatumomab versus standard care of £28,655 per QALY gained.

The ERG critically appraised the company's economic analysis and double-programmed the deterministic version of the company's model. The ERG's critical appraisal identified several issues relating to the company's economic analysis and the evidence used to inform it. These include: (i) the exclusion of relevant patient subgroups from the model; (ii) the exclusion of the "monitor for relapse" comparator from the analysis; (iii) use of a model structure which is inappropriate for tracking HSCT; (iv) the absence of RCT evidence for blinatumomab versus standard care; (v) concerns regarding the company's approach to RFS/OS model selection; (vi) concerns regarding the robustness of the

company's alternative base case (blinatumomab used on relapse for the standard care group); (vii) the questionable reliability of the company's HRQoL estimates; (viii) uncertainty surrounding the proportion of RFS events that are deaths; (ix) the inclusion of an unrealistic treatment pathway and (x) limited sensitivity analysis around alternative parametric functions.

The ERG undertook eight sets of exploratory analyses using the deterministic version of the company's updated model. Notwithstanding uncertainty relating to the choice of parametric RFS and OS functions, the ERG's preferred model includes the correction of seven minor programming errors and the inclusion of a fixed 5-year cure point. The ERG-preferred model produces a deterministic ICER for blinatumomab versus standard care of £30,227 per QALY gained. The ERG also undertook a number of further analyses to explore the impact of alternative parametric models and alternative parameter values on the model results. These analyses indicate that the costs of standard chemotherapy, the post-HSCT survival probabilities and the utility value for the post-relapse state have only a minor impact on the ICER for blinatumomab versus standard care. Conversely, the cure fraction and the choice of parametric OS distribution have a significant impact on the ICER for blinatumomab versus standard care. Within the ERG's exploratory analysis of alternative RFS and OS functions, the ICER for blinatumomab versus standard care ranges from £25,783 per QALY gained (Weibull non-mixture cure model, unrestricted) to £63,265 per QALY gained (Weibull, unrestricted). Across the full range of models considered, only the Weibull non-mixture cure model (unrestricted) and the Weibull mixture cure model (unrestricted) produce results in which the full range of ICERs are below £30,000 per QALY gained (irrespective of RFS model assumed). The clinical advisors' three preferred OS models (Generalised gamma [unrestricted], RCS Weibull [unrestricted] and Weibull mixture cure [unrestricted]) result in ICERs in the range £25,810 per QALY gained to £34,904 per QALY gained.

The ERG considers the following to represent the key uncertainties within the company's health economic analysis:

- The absence of comparative clinical and economic evidence for blinatumomab versus standard care chemotherapy within subgroups of BLAST which were excluded from the comparative analysis (patients with Ph+ MRD+ BCP-ALL and patients with Ph- MRD+ BCP-ALL with CR2+).
- The absence of clinical data and economic comparisons of blinatumomab versus monitoring for patients who are unable to undergo HSCT or to tolerate chemotherapy.
- The necessary reliance on adjusted historical control evidence, due to the absence of RCT evidence for blinatumomab versus standard care, and the potential for unobserved confounders.
- The long-term extrapolation of RFS and OS outcomes, including the timing of cure.

6. END OF LIFE

NICE end of life supplementary advice should be applied in the following circumstances and when both the criteria referred to below are satisfied:

- The treatment is indicated for patients with a short life expectancy, normally less than 24 months and;
- There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared to current NHS treatment.

The CS[1] states that blinatumomab meets NICE's criteria for life-extending therapies given at the end of life. The company's evidence supporting this is presented in Table 51.

Table 51: Evidence supporting the company's end of life argument (reproduced from CS Table 50)

Criterion	Data available
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	Median OS, using ATT-weighted propensity score matching analyses for standard care chemotherapy was [redacted]. The estimated mean survival (undiscounted) in the economic analysis was almost 5x greater than the median survival (7.86 years) in the standard care arm; however, this is reflective of the small proportion of patients who achieve long-term survival (~20%). For this reason, the median survival is considered to be a more suitable representation of the anticipated survival in the patient population as a whole.
There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment	Median OS, using ATT-weighted propensity score matching analyses (Section B.2.9.4), was [redacted] after more than 40 months follow-up for blinatumomab thus demonstrating a [redacted] OS survival [redacted] when compared to standard care. The estimated mean survival (undiscounted) in the economic analysis was [redacted] years in the blinatumomab arm, resulting in an incremental survival benefit of [redacted] years.

The CS argues that a small number of patients in the historical control study who received standard care chemotherapy were observed to survive for a long time and that *"[Due to] the skew caused by this small group of patients, it was considered appropriate to use median OS values, rather than the mean, so as to more accurately represent the patient population as a whole. This skew effect and use of median OS rather than the mean has been noted in previous appraisals where the Committee agreed that consideration of medians was more appropriate"* (CS,[1] page 84).

The ERG strongly disagrees with the company's proposed use of median values to determine whether NICE's end of life criteria are met. Medians represent the "middle patient" and do not take account of

skewness in the distribution of patient outcomes; conversely, the only measure of central tendency which fully represents outcomes for the population as a whole is the mean. Given the use of parametric cure models to inform OS, the mean and median OS estimates generated by the company's model diverge significantly (blinatumomab median OS=5.85 years versus mean [undiscounted] OS=13.59 years; standard care median OS=1.86 versus mean [undiscounted] OS=7.86 years). Based on the ERG's exploratory analyses, the lowest (undiscounted) mean OS for the standard care group across all models considered is 7.69 life years; all OS models suggest an undiscounted incremental OS gain of 2.12 years or greater. On the basis of these exploratory analyses, the ERG does not believe that blinatumomab meets NICE's criteria for life-extending treatments given at the end of life. The ERG also notes that due to the absence of a head-to-head RCT comparing blinatumomab against a relevant comparator, and the necessary use of a statistical matching approach to inform indirect treatment comparisons, there is uncertainty surrounding the true magnitude of OS benefit attributable to blinatumomab.

7. OVERALL CONCLUSIONS

In the absence of direct comparative data with other treatments, the main evidence in the CS was derived from two single-arm open-label studies of blinatumomab, of which one was a pilot study which was not used for the comparison with standard care chemotherapy. The two blinatumomab studies were well conducted, however single-arm studies are subject to biases. The main blinatumomab evidence came from the BLAST study of 116 patients. One historical control study (Study 20120148) of standard care chemotherapy was included (n=287); this study that analysed data from existing clinical databases.

From the 116 patients in BLAST, median OS was [redacted], with an OS at 18 months follow-up of [redacted]. From 110 patients providing RFS data from BLAST, median RFS was [redacted]; RFS at 18 months was [redacted]. Based on the EORTC QLQ-C30, patients reported [redacted]. By the end of the core study, [redacted] HRQoL as measured by EQ-5D did not change significantly from baseline to the end of the core study. [redacted] participants experienced at least one treatment-emergent AE. Comparative effectiveness for patients with Ph- MRD+ BCP-ALL in CR1 was estimated through indirect comparison of the BLAST PAS data and a historical control study using ATT propensity score weights. This analysis suggested an HR [redacted].

Notwithstanding uncertainty relating to the choice of parametric RFS and OS functions, the ERG's preferred analysis increases the ICER for blinatumomab versus standard care from £27,779 to £30,227 per QALY gained; this difference is driven by the inclusion of a structural cure assumption for surviving patients at 5-years. Additional exploratory undertaken by the ERG suggests that that the costs of standard care chemotherapy, the post-HSCT survival probabilities and the utility value for the post-relapse state have only a minor impact on the ICER for blinatumomab versus standard care. Conversely, the cure fraction and the choice of parametric OS distribution have a significant impact on the ICER for blinatumomab versus standard care. Within the ERG's exploratory analysis of alternative RFS and OS functions, the ICER for blinatumomab versus standard care ranges from a lowest ICER of £25,783 per QALY gained (unrestricted Weibull non-mixture cure model) to a highest ICER of £63,265 per QALY gained (unrestricted Weibull). Across the full range of models considered, only the Weibull non-mixture cure model (unrestricted) and the Weibull mixture cure model (unrestricted) produce results in which the full range of ICERs are below £30,000 per QALY gained (irrespective of RFS model assumed). The clinical advisors' three preferred OS models (Generalised gamma [unrestricted], RCS Weibull [unrestricted] and Weibull mixture cure [unrestricted]) result in ICERs in the range £25,810 per QALY gained to £34,904 per QALY gained. The ERG notes that all analyses should be considered

highly uncertain due to the absence of RCT evidence for blinatumomab versus standard care and a lack of evidence relating to long-term RFS and OS outcomes for patients treated with blinatumomab (including the timing of cure). The ERG further notes that no comparative clinical or economic evidence is available for the comparison of blinatumomab versus standard care chemotherapy in patients Ph+ MRD+ BCP-ALL or in MRD+ BCP-ALL patients in CR2+, or for the comparison of blinatumomab versus monitoring in patients who are unable to undergo HSCT or to tolerate chemotherapy.

The ERG does not believe that blinatumomab meets NICE's criteria for life-extending treatments given at the end of life.

8. REFERENCES

1. Amgen Ltd. Blinatumomab for acute lymphoblastic leukaemia for people with minimal residual disease activity in remission [ID1036]. Company's evidence submission to the National Institute for Health and Care Excellence. In. Cambridge, UK: Amgen; 2017.
2. Alvarnas JC, Brown PA, Aoun P, Ballen KK, Bellam N, Blum W, *et al.* Acute lymphoblastic leukemia. *J Natl Compr Canc Netw* 2012;10:858-914.
3. National Comprehensive Cancer Network. NCCN Clinical Practice Guidelines in Oncology: Acute Lymphoblastic Leukemia. Version 1.2014. 2014. https://www.nccn.org/professionals/physician_gls/f_guidelines.asp (Accessed 23 November 2017).
4. Goekbuget N, Kneba M, Raff T, Trautmann H, Bartram CR, Arnold R, *et al.* Adult patients with acute lymphoblastic leukemia and molecular failure display a poor prognosis and are candidates for stem cell transplantation and targeted therapies. *Blood* 2012;120:1868-76.
5. Thomas X, Boiron J-M, Huguet F, Dombret H, Bradstock K, Vey N, *et al.* Outcome of treatment in adults with acute lymphoblastic leukemia: analysis of the LALA-94 trial. *Journal of clinical oncology* 2004;22:4075-86.
6. Moorman AV, Harrison CJ, Buck GA, Richards SM, Secker-Walker LM, Martineau M, *et al.* Karyotype is an independent prognostic factor in adult acute lymphoblastic leukemia (ALL): analysis of cytogenetic data from patients treated on the Medical Research Council (MRC) UKALLXII/Eastern Cooperative Oncology Group (ECOG) 2993 trial. *Blood* 2007;109:3189-97.
7. Oriol A, Vives S, Hernández-Rivas J-M, Tormo M, Heras I, Rivas C, *et al.* Outcome after relapse of acute lymphoblastic leukemia in adult patients included in four consecutive risk-adapted trials by the PETHEMA Study Group. *Haematologica* 2010;95:589-96.
8. Chiaretti S, Vitale A, Cazzaniga G, Orlando SM, Silvestri D, Fazi P, *et al.* Clinico-biologic features of 5202 acute lymphoblastic leukemia patients enrolled in the Italian AIEOP and GIMEMA Protocols and stratified in age-cohorts. *Haematologica* 2013:haematol. 2012.080432.
9. Fielding AK, Richards SM, Chopra R, Lazarus HM, Litzow MR, Buck G, *et al.* Outcome of 609 adults after relapse of acute lymphoblastic leukemia (ALL); an MRC UKALL12/ECOG 2993 study. *Blood* 2007;109:944-50.
10. Moorman AV, Chilton L, Wilkinson J, Ensor HM, Bown N, Proctor SJ. A population-based cytogenetic study of adults with acute lymphoblastic leukemia. *Blood* 2010;115:206-14.
11. Radich JP. Philadelphia chromosome-positive acute lymphocytic leukemia. *Hematol Oncol Clin North Am* 2001;15:21-36.
12. Fielding AK. Current treatment of Philadelphia chromosome-positive acute lymphoblastic leukemia. *Hematology Am Soc Hematol Educ Program* 2011:231-7.
13. Hoelzer D, Bassan R, Dombret H, Fielding A, Ribera J, Buske C. Acute lymphoblastic leukaemia in adult patients: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology* 2016;27:v69-v82.
14. Campana D. Minimal residual disease in acute lymphoblastic leukemia. *ASH Education Program Book* 2010;2010:7-12.
15. Bassan R, Spinelli O, Oldani E, Intermesoli T, Tosi M, Peruta B, *et al.* Improved risk classification for risk-specific therapy based on the molecular study of minimal residual disease (MRD) in adult acute lymphoblastic leukemia (ALL). *Blood* 2009;113:4153-62.

16. Holowiecki J, Krawczyk-Kulis M, Giebel S, Jagoda K, Stella-Holowiecka B, Piatkowska-Jakubas B*, et al.* Status of minimal residual disease after induction predicts outcome in both standard and high-risk Ph-negative adult acute lymphoblastic leukaemia. The Polish Adult Leukemia Group ALL 4-2002 MRD Study. *British Journal of Haematology* 2008;142:227-37.

17. Amgen Ltd. Blinatumomab for acute lymphoblastic leukaemia for people with minimal residual disease activity in remission. Response to clarification questions. In. Cambridge, UK: Amgen; 2017.

18. 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:e170580.

19. UKALL14. A randomized trial for adults with newly diagnosed acute lymphoblastic leukaemia. 2012. 201210/01/2018).

20. Bruggemann M, Raff T, Flohr T, Goekbuget N, Nakao M, Droese J*, et al.* Clinical significance of minimal residual disease quantification in adult patients with standard-risk acute lymphoblastic leukemia. *Blood* 2006;107:1116-23.

21. Amgen Inc. Data on File: Minimal residual disease testing patterns and healthcare utilization for adults with B-precursor acute lymphoblastic leukemia - results of a physician survey in the UK. In; 2017.

22. National Institute for Health and Care Excellence. Blinatumomab for acute lymphoblastic leukaemia for people with minimal residual disease activity in remission: final scope. In. London, UK.: NICE; 2017.

23. van Dongen JJ, van der Velden VH, Brüggemann M, Orfao A. Minimal residual disease diagnostics in acute lymphoblastic leukemia: need for sensitive, fast, and standardized technologies. *Blood* 2015;125:3996-4009.

24. European Medicines Agency. Summary of Product Characteristics - blinatumomab. EMA. https://www.medicines.org.uk/emc/product/5064 (Accessed 12/01/2018).

25. Royal Pharmaceutical Society. Medicines Complete (BNF). 2017.

26. National institute for Health and Care Excellence. Guide to the methods of technology appraisal. London, UK: NICE; 2013.

27. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011] edn: The Cochrane Collaboration; 2011.

28. Sterne JAC, 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.

29. Dalziel K, Round A, Stein K, Garside R, Castelnuovo E, Payne L. Do the findings of case series studies vary significantly according to methodological characteristics? *Health technology assessment (Winchester, England)* 2005;9:iii-iv, 1-146.

30. Centre for Reviews and Dissemination. Systematic Reviews: CRD's Guidance for Undertaking Reviews in Healthcare. 3rd Edition edn. York: CRD, University of York; 2009.

31. Amgen Inc. Clinical Study Report: MT103-203 Addendum 1 (Key Secondary Analysis). In; 2016.

32. Amgen Inc. Clinical Study Report: MT103-202. In. Cambridge, UK; 2013.

33. Amgen Inc. Observational Research Study Report: A Retrospective Analysis of Hematological Relapse Free Survival (RFS) and Overall Survival (OS) in Adult Patients with Philadelphia-Negative B-Precursor Acute Lymphoblastic Leukemia (ALL) in Complete Hematological Remission with Minimal Residual Disease (MRD). . In; 2017.

34. Evans SR. Clinical trial structures. *Journal of experimental stroke & translational medicine* 2010;3:8-18.

35. Chambers D, Rodgers M, Woolacott N. Not only randomized controlled trials, but also case series should be considered in systematic reviews of rapidly developing technologies. *Journal of clinical epidemiology* 2009;62:1253-60.e4.

36. Jüni P, Altman DG, Egger M. Assessing the quality of controlled clinical trials. *BMJ* 2001;323:42-6.

37. Sedgwick P. Retrospective cohort studies: advantages and disadvantages. *BMJ : British Medical Journal* 2014;348.

38. Amgen Inc. Study Protocol: A confirmatory multicenter, single-arm study to assess the efficacy, safety, and tolerability of the BiTE antibody blinatumomab in adult patients with minimal residual disease (MRD) of B-precursor acute lymphoblastic leukemia. In; 2010.

39. Goekbuget N, Dombret H, Bonifacio M, Reichle A, Graux C, Havelange V, *et al.* BLAST: A Confirmatory, Single-Arm, Phase 2 Study of Blinatumomab, a Bispecific T-Cell Engager (BiTE®) Antibody Construct, in Patients with Minimal Residual Disease B-Precursor Acute Lymphoblastic Leukemia (ALL). *Blood* 2014;124:379.

40. Gökbuget N, Zugmaier G, Klinger M, Kufer P, Stelljes M, Viardot A, *et al.* Long-term relapse-free survival in a phase 2 study of blinatumomab for the treatment of patients with minimal residual disease in B-lineage acute lymphoblastic leukemia. *Haematologica* 2017;102:e132-e5.

41. Pulte D, Jansen L, Castro FA, Brenner H. Changes in the survival of older patients with hematologic malignancies in the early 21st century. *Cancer* 2016;122:2031-40.

42. Fayers PM AN, Bjordal K, et al. . . The EORTC QLQ-C30 Scoring Manual (3rd Edition). 2001.

43. European Medicines Agency. EPAR for Blincyto (blinatumomab). http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/003731/human_med_001921.jsp&mid=WC0b01ac058001d124 (Accessed 28th January).

44. Amgen Inc. 203-148-MRD-propensity-score-analysis-report-att-updated-final. In; 2017.

45. Faria R, Hernandez Alava M, Manca A, Wailoo A. NICE DSU technical support document 17: the use of observational data to inform estimates of treatment effectiveness for technology appraisal: methods for comparative individual patient data. *Sheffield: NICE Decision Support Unit* 2015.

46. Cole SR, Hernan MA. Adjusted survival curves with inverse probability weights. *Computer methods and programs in biomedicine* 2004;75:45-9.

47. National institute for Health and Care Excellence. Inotuzumab ozogamicin for treating relapsed or refractory B-cell acute lymphoblastic leukaemia. London; 2017.

48. UK Stem Cell Strategy Oversight Committee. NHS Blood and Transplant. Unrelated Donor Stem Cell Transplantation in the UK. Effective Affordable Sustainable; 2014.

49. van Agthoven M, Groot MT, Verdonck LF, Lowenberg B, Schattenberg AV, Oudshoorn M, *et al.* Cost analysis of HLA-identical sibling and voluntary unrelated allogeneic bone marrow and peripheral blood stem cell transplantation in adults with acute myelocytic leukaemia or acute lymphoblastic leukaemia. *Bone marrow transplantation* 2002;30:243-51.

50. Kind P, Hardman G, Macran S. UK Population Norms for EQ-5D. University of York; 1999.

51. Kurosawa S, Yamaguchi T, Mori T, Kanamori H, Onishi Y, Emi N, *et al.* Patient-reported quality of life after allogeneic hematopoietic cell transplantation or chemotherapy for acute leukemia. *Bone marrow transplantation* 2015;50:1241-9.

52. Commercial Medicines Unit. Drugs and Pharamaceutical electronic market information (eMit). 2017. https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit (Accessed 11/12/2017).

53. National Institute for Health and Care Excellence. Blinatumomab for treating Philadelphiachromosome-negative relapsed or refractory acute lymphoblastic leukaemia. London, UK: NICE; 2017.

54. Addicott R, Dewar S. Improving choice at end of life: a descriptive analysis of the impact and costs of the Marie Cure Delivering Choice Programme in Lincolnshire. . London, UK: King's Fund; 2008.

55. Marie Curie Cancer Care. Understanding the cost of end of life care in different settings. 2012. https://www.mariecurie.org.uk/globalassets/media/documents/commissioning-our-services/publications/understanding-cost-end-life-care-different-settingspdf (Accessed 11/01/2018).

56. Department of Health. NHS Reference Costs 2015 to 2016. 2016. https://www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016 (Accessed 10/01/2018).

57. Collett D. Modelling survival data in medical research: CRC press; 2015.

58. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. *Value in Health* 2010;13:509-18.

59. Drummond M, Sculpher M, Torrance G, O'Brien B, Stoddart G. Methods for the Economic Evaluation of Health Care Programmes. 3rd edn. New York: Oxford University Press; 2005.

60. Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C, *et al.* Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices--Modeling Studies. *Value in Health* 2003;6:9-17.

61. Amgen Inc. ALL epidemiology estimates. *Company evidence submission template for blinatumomab for acute lymphoblastic leukaemia for people with minimal residual disease activity in remission [ID1036],* 2017.

62. Goldstone AH, Richards SM, Lazarus HM, Tallman MS, Buck G, Fielding AK, *et al.* In adults with standard-risk acute lymphoblastic leukemia, the greatest benefit is achieved from a matched sibling allogeneic transplantation in first complete remission, and an autologous transplantation is less effective than conventional consolidation/maintenance chemotherapy in all patients: final results of the International ALL Trial (MRC UKALL XII/ECOG E2993). *Blood* 2008;111:1827-33.

9. APPENDICES

Appendix 1: Technical appendix detailing methods for applying the ERG's exploratory analyses within the company's model

Exploratory analysis 1 – correction of model errors

ERG exploratory analysis 1 corrects five errors which were not previously addressed within the company's updated model.

1. Annual general population mortality rate applied for 1-year intervals defined according to time since model entry, rather than according to patient age. Not corrected in company's updated model.
2. Risk of all-cause death exceeds 1.0 for males patients aged 95 years and older and female patients aged 97 years and older. Corrected in company's updated model.
3. Minor programming errors in formula used to calculate receipt of HSCT at 2 years. Not corrected in company's updated model.
4. Incorrect formula offset in discounting cost of other inpatient visits post-relapse in the blinatumomab group. Not corrected in company's updated model.
5. Post-relapse HSCTs assumed to occur after the 5-year time point; inconsistent with the rest of the model structure. Corrected in company's updated model.
6. Inappropriate application of utility decrement due to proximity to death. Not corrected in company's updated model.
7. Incorrectly discounting of HSCT costs due to the use of approximate =LOOKUP() functions used to calculate the discount rate for receipt of HSCT. Not corrected in company's updated model.

1. Correct mortality lookup error

a. Open the model ID1026 Blin MRD Ph- B ALL_Updated CEM.xlsb
b. Go to "Blin Calc" worksheet
c. Go to cell AW9
d. Type the formula "=VLOOKUP(ROUNDDOWN(AT9,0),$AL$9:$AR$91,4,TRUE)"
e. Copy the formula down column AW
f. Go to cell AX9
g. Type the formula "=MAX(VLOOKUP(ROUNDDOWN(AT9,0)+1,$AL$9:$AR$91,4,TRUE),0)"
h. Copy the formula down column AX
i. Go to "SOC Calc" worksheet
j. Go to cell AW9
k. Type the formula "=VLOOKUP(ROUNDDOWN(AT9,0),'SOC Calc'!$AL$9:$AP$91,4,TRUE)"
l. Copy the formula down column AW
m. Go to cell AX9

n. Type the formula "=MAX(VLOOKUP(ROUNDDOWN(AT9,0)+1,'SOC Calc'!$AL$9:$AP$91,4,TRUE),0)"
o. Copy the formula down column AX9

3. HSCT programming error

a. Go to "Blin Calc" worksheet
b. Go to cell GZ13
c. Type the formula "=GY12*(hsct.pctmo24/hsct.pctmo13)"
d. Copy the formula down column GZ
e. Go to cell HQ13
f. Type the formula "=HP12*(hsct.pctmo24/hsct.pctmo13)"
g. Copy the formula down column HQ
h. Go to "SOC Calc" worksheet
i. Go to cell GZ13
j. Type the formula" =GY12*(hsct.pctmo24/hsct.pctmo13)"
k. Copy the formula down column GZ
l. Go to cell HQ 13
m. Type the formula "=HP12*(hsct.pctmo24/hsct.pctmo13)"
n. Copy the formula down column HP

4. Correct cost discounting formula for blinatumomab

a. Go to "Blin Calc" worksheet
b. Go to cell E27
c. Type the formula "=SUMPRODUCT($GC$9:$GC$3138,$M$9:$M$3138,R9:R3138)"

6. Correction proximity to death decrement

a. Go to "Blin Calc" worksheet
b. Go to cell JH9
c. Type the formula "=IF(IO9<model.term_util_end,MIN(J9,0.5),0)"
d. Copy the formula down column JH
e. Go to cell JI9
f. Type the formula "=JG9*util.term*JH9"
g. Copy the formula down column JI
h. Go to "SOC Calc" worksheet
i. Go to cell JH9
j. Type the formula "=IF(IO9<model.term_util_end,MIN(J9,0.5),0)"
k. Copy the formula down column JH
l. Go to cell JI9
m. Type the formula "=JG9*util.term*JH9"
n. Copy the formula down column JI

7. Incorrectly discounting of HSCT costs

a. Go to "Blin Calc" worksheet
b. Go to cell GS9
c. Type the formula "=1/(1+_input_model.discount_cost)^ROUNDDOWN(GQ9,0)"
d. Copy the formula down column GS
e. Go to the "SOC Calc" worksheet

f. Go to cell GS9
g. Type the formula "=1/(1+_input_model.discount_cost)^ROUNDDOWN(GQ9,0)"
h. Copy the formula down column GS

Exploratory analysis 2 – application of cure point at 5-years

a. Go to worksheet "Settings" cell G26.
b. Select "switch" from the drop-down menu.

Exploratory analysis 3 – ERG-preferred analysis

Combine exploratory analysis 1 and 2.

All subsequent exploratory analyses are based on this version of the model.

Exploratory analysis 4 - impact of alternative standard care chemotherapy costs

a. Go to the "Cost Inputs" sheet
b. Go to cell F84
c. Type the formula "=29.26*2"
d. Go to cell G84
e. Type the formula "=0.41*2"
f. Go to cell H84
g. Type the formula "=49.15*2"
h. Go to cell I84
i. Type the formula "=4.39*2"
j. Go to cell J84
k. Type the formula "=6.63*2"

Exploratory analysis 5 - alternative post-HSCT survival probabilities

a. Go to the "Blin calc" worksheet
b. Insert a new column GR
c. Go to cell GR8 type Age(years)
d. Go to cell GR9 use the formula "=ROUNDDOWN($K$9+GQ9,0)"
e. Copy this formula down column GR
f. Insert three new columns GS, GT, and GU
g. Label column GS " probability of death (1st 6 months)"
h. Label Column GT probability of death (future months)
i. Label column GU rate of death
j. Go to cell AS8 type "Gender weighted probability of dying between ages"
k. Go to cell AS9 and type the formula "=model.pct_male*AM9+(1-model.pct_male)*AN9"
l. Copy this formula down column AS
m. Go to cell GU and type the following formula "=-(LN(1-VLOOKUP(GR9,$AL$9:$AS$91,8,FALSE))/365.25)"
n. Copy this formula down GU
o. Go to cell GS9 and type the formula =IFERROR(0.07+1-EXP(-GU9*((365.25/2)-100)),100%)"
p. Copy this formula down column GS
q. Go to cell GT9 and type the formula "=IFERROR(1-EXP(-GU9*((365.25/2))),100%)"
r. Copy this formula down column GT
s. Go to Cell GZ9 and type the formula "=GY9*(1-$GS9)"
t. Copy this formula down column GZ

u. Go to cell HA10 type the formula "=GZ9*(1-$GT10)"
v. Copy down
w. Copy cell HA10
x. Paste the formula into cells HB11, HC12, HD13, HE14, HF15, HG16, HH17, HI18
y. Copy the formulae down columns HA, HB, HC, HD, HE, HF, HG, HH
z. Go to cell HJ19 and type the formula " =(HI18+HJ18)*(1-$GT19)"
aa. Copy down column HJ
bb. Select cells GZ9:HJ129
cc. Copy the cells
dd. Select cell HQ9
ee. Paste the formulae
ff. Go to the SOC Calc worksheet
gg. Go to cell GV9 and type the formula "=GU9*(1-'Blin Calc'!$GS9)"
hh. Copy this formula down column GV.
ii. Go to cell GW10 and type the formula "=GV9*(1-'Blin Calc'!$GT10)"
jj. Copy this formula down column GW
kk. Copy cell GW10
ll. Paste the formula into cells GX11, GY12, GZ13, HA14, HB15, HC16, HD17, HE18
mm. Copy down columns GX, GY, GZ, HA, HB, HC, HD, HE
nn. Go to cell HF19 and type "=(HE18+HF18)*(1-'Blin Calc'!$GT19)"
oo. Copy this formula down column HF
pp. Copy cells GV9:HF129
qq. Select cell HM9
rr. Paste the formulae

Exploratory analysis 6 - alternative cure fractions for the standard care group

a. Go to worksheet "SOC Calc" cell CM15
b. Apply alternative cure fractions

Exploratory analysis 7 - alternative post-relapse utilities

a. Go to worksheet "Utility Inputs" cell F18
b. Apply alternative post-relapse utility values

Exploratory analysis 8 - Exploration of the impact of alternative parametric RFS and OS models

Run macro as per instructions provided by the company using ERG-preferred model

Appendix 2: Blinatumomab for acute lymphoblastic leukaemia for people with minimal residual disease activity in remission - Model selection exercise

Background information

Within their health economic model, the company has fitted a range of parametric survivor functions to time-to-event outcomes (overall survival [OS] and relapse-free survival [RFS]) for patients with Ph-disease with CR1 from the BLAST study and the ATT-weighted historical control study in order to extrapolate beyond the duration of the empirical studies. These survival curves influence both the costs and the health gains predicted by the company's model. We have some concerns regarding how the company has selected their preferred survival curves for use in the model, particularly with respect to the plausibility of the extrapolated portion of the curve. Our main concern is surrounding OS, as this is a key driver of the cost-effectiveness of blinatumomab.

Based on clinical advice, we believe that it would be broadly appropriate to assume that patients who have not relapsed within 5-years are cured. For simplicity, we have assumed the same to be true with respect to OS, although we note that some relapsed patients may achieve cure as a consequence of downstream treatments (e.g. HSCT received post-relapse), hence the time at which cure manifests may be slightly later for OS than RFS.

We have plotted the Kaplan-Meier curves from the BLAST and historical control studies and have overlaid these with a range of long-term potential OS survivor functions (see Figures 1 and 2). As a consequence of the assumption of cure at 5-years, all models are based on the company's statistical model projections for up to 5-years; the survivor function is then applied using uplifted general population mortality rates thereafter. The model assumes a population starting age of roughly 45 years.

Your task

We now need to choose which curve is likely to be most appropriate for OS. We would like you to look at the fitted curves presented in Figure 1 (blinatumomab OS) and Figure 2 (standard care OS) and to fill in the responses to questions on pages 4 and 5 to indicate which of the curves you consider to be the most clinically plausible and to state your reasons why. In completing this exercise, please consider both how well the curve appears to fit the observed data as well as the clinical plausibility of the extrapolation beyond the observed period. To do this you may wish to think about:

- The distance between the smooth parametric curves and the stepped Kaplan-Meier function (note that the end of the Kaplan-Meier curve is very uncertain)
- The proportion of patients you would expect to achieve a cure by 5-years
- The probability of surviving at different timepoints in each treatment group

We note that several of the curves appear to be very similar. If you wish to select multiple preferred curves, please do so.

Figure 1: Comparison of alternative OS survivor functions (including 5-year cure assumption) - blinatumomab

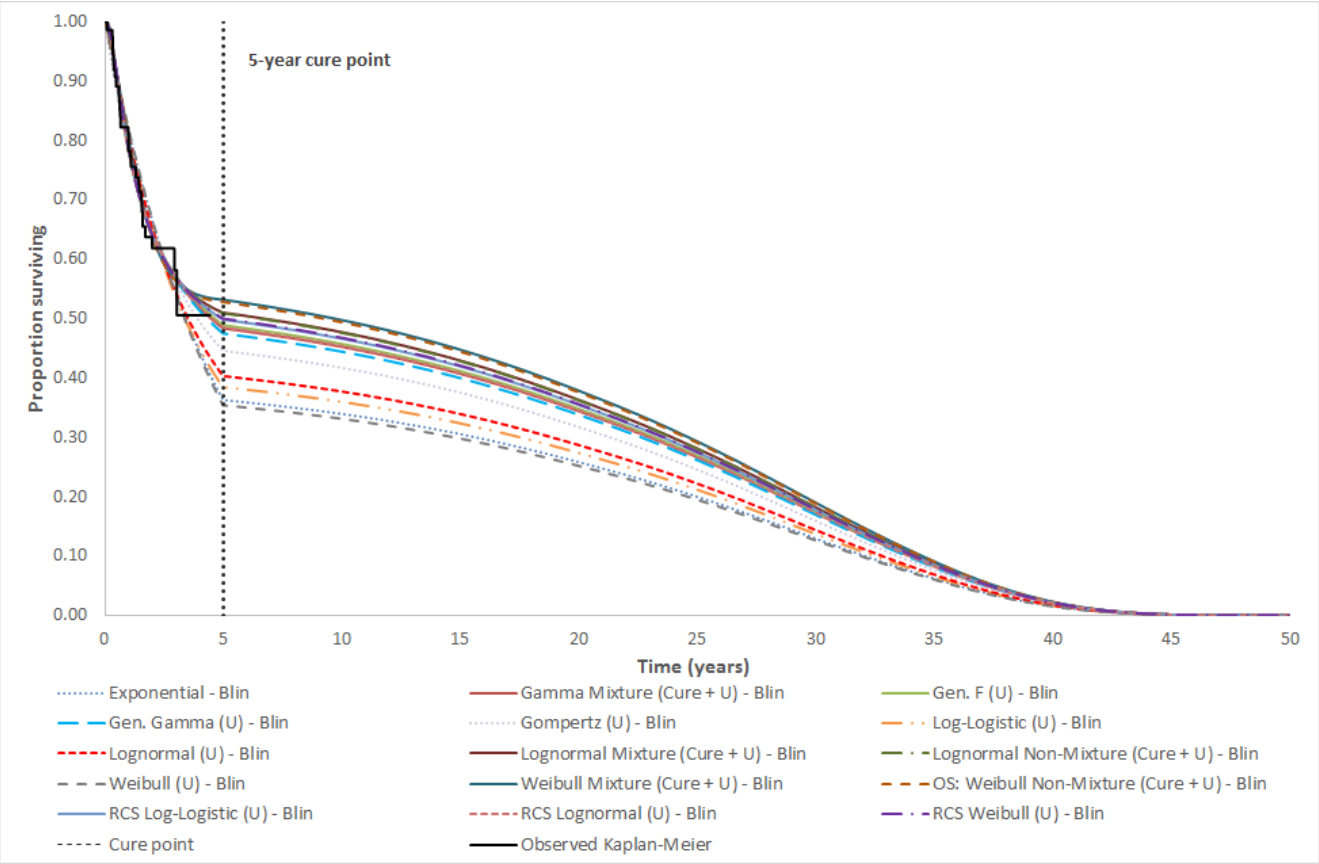
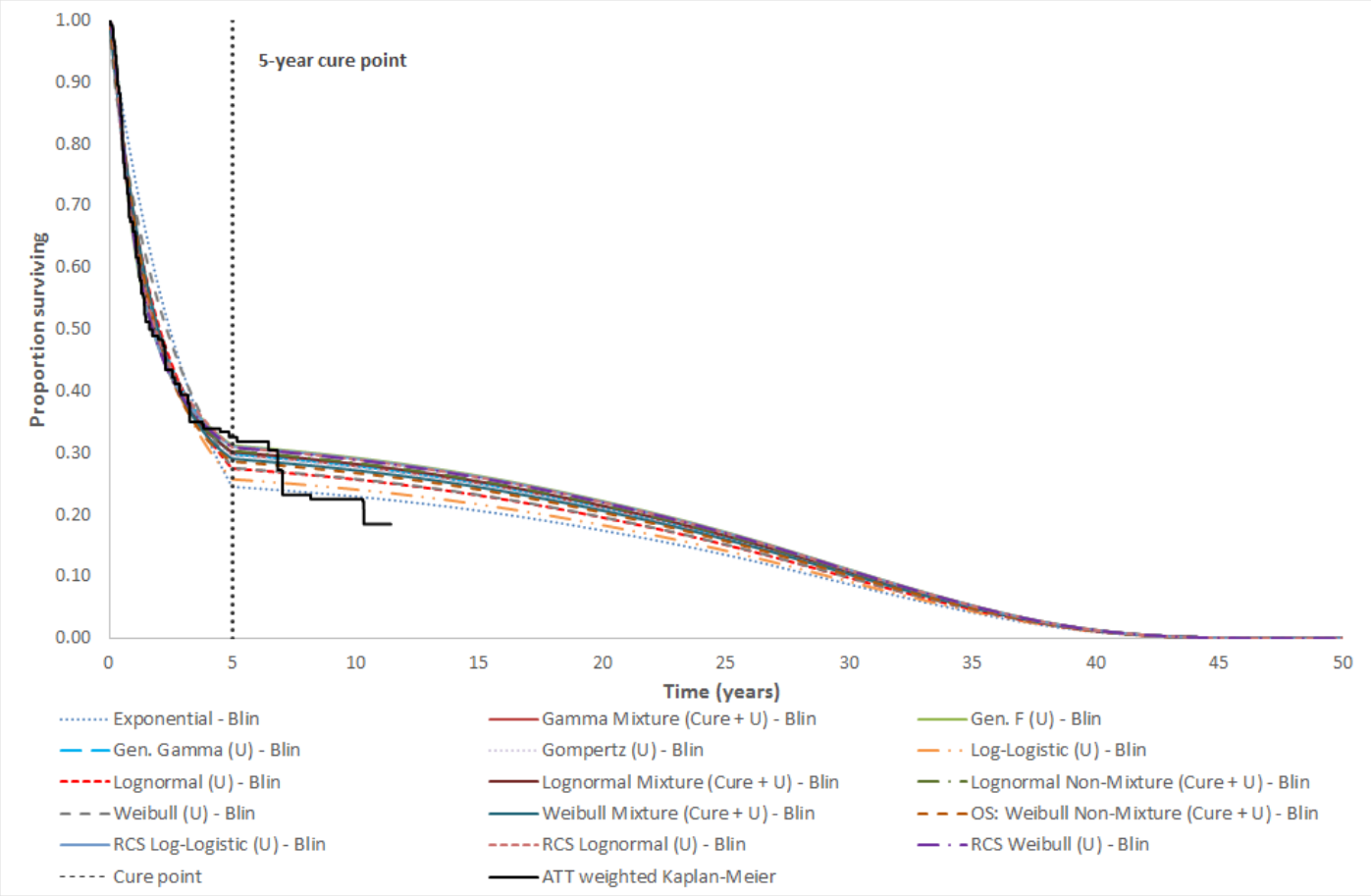


Figure 2: Comparison of alternative OS survivor functions (including 5-year cure assumption) – standard care



Clinicians' responses

QUESTION 1. Do you think it is reasonable to apply the cure point at 5-years? Or should we assume a later timepoint for OS?

RESPONSE 1:

Blinatumomab group (Please refer to Figure 1)

QUESTION 2. Which is your preferred OS function for the blinatumomab group?

RESPONSE 2:

QUESTION 3. Please state why this is your preferred function

RESPONSE 3:

QUESTION 4. Which other functions would you consider to be plausible?

RESPONSE 4:

Standard care group (Please refer to Figure 2)

QUESTION 5. Which is your preferred OS function for the standard care group?

RESPONSE 5:

_____

QUESTION 6. Please state why this is your preferred function

RESPONSE 6:

_____

QUESTION 7. Which other functions would you consider to be plausible?

RESPONSE 7:

_____

National Institute for Health and Care Excellence
Centre for Health Technology Evaluation

Pro-forma Response

ERG report

Blinatumomab for treating minimal residual B-precursor acute lymphoblastic leukaemia in remission [ID1036]

You are asked to check the ERG report from ScHARR to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies you must inform NICE by **5pm on Monday 19 February 2018** using the below proforma comments table. All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the Evaluation report.

The proforma document should act as a method of detailing any inaccuracies found and how and why they should be corrected.

Thank you for the opportunity to review the report for factual inaccuracies. In general, we found the report to be fair, thorough and of a high-quality. Our response below is structured in two parts:

- ***Section 1**: Issues pertaining to the ERG's interpretation of our evidence submission that lead to misleading and potentially inaccurate conclusions*
- ***Section 2**: Factual inaccuracies (eg. typographical or reporting mistakes and confidentiality marking).*

Issue 1 'Monitor for relapse' as a comparator

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 3 Sentence: 'The CS also excludes the comparator of "monitor for relapse" based on the argument that it is highly unlikely that MRD+ patients who are at high risk of relapse would not receive active treatment. However, clinical advisors to the Evidence Review Group (ERG) noted that due to its favourable toxicity profile, blinatumomab may be a potential treatment option for patients who are unable to undergo haematopoietic stem cell transplantation (HSCT) or to tolerate chemotherapy; the ERG considers that a further comparison of blinatumomab versus monitoring within this subgroup should have been explored.' This critique should include the context, provided later on p. 22, that the clinical advisors to the ERG noted that 'this population [of older and less fit patients] would be small'. This is aligned with the justification provided in our submission and the feedback heard from clinical advisors during initial scoping discussions –	Amend to: 'The CS also excludes the comparator of "monitor for relapse" based on the argument that it is highly unlikely that MRD+ patients who are at high risk of relapse would not receive active treatment. Clinical advisors to the Evidence Review Group (ERG) noted that due to its favourable toxicity profile, blinatumomab may be a potential treatment option for patients who are unable to undergo haematopoietic stem cell transplantation (HSCT) or to tolerate chemotherapy; the ERG considers that a further comparison of blinatumomab versus monitoring within this subgroup should have been explored. **However, the clinical experts also noted that this subgroup is likely to be small and it is unclear whether any relevant comparator data exist**.''	The ERG report unfairly criticises the submission for excluding a subgroup that is deemed minor and which has limited available evidence to inform a comparison.	This is not a factual inaccuracy. The company ruled out monitor for relapse in the decision problem section and did not consider this comparator further within the submission. The ERG believes that the company should have explored this comparison, irrespective of subgroup size, as it is within the remit of the final NICE scope. Statements regarding the availability of evidence should have been made by the company based on robust systematic review methods, rather than by the ERG, based on expectations. No amendment has been made.

specifically, that due to the high risk of relapse in this this disease area it is highly unlikely that patients would not receive active treatment. In addition, on p. 4 and p. 57 the ERG notes that 'it is unclear whether any relevant comparator data exist within this subgroup' and that 'it is unlikely that any relevant studies of blinatumomab […] have been missed'. These statements provide justification for the absence of this minor subgroup from the company submission. Furthermore, BLAST eligibility criteria does not explicitly require patients to be transplant eligible and/or able to tolerate chemotherapy and the historical cohort study includes a cohort of patients receiving SOC, which may include both active treatment and monitoring for relapse (without treatment). However, it was not possible to determine the percentage of patients in the historical comparator who did not receive active treatment as this information was not collected. Similar statements are repeated on p. 4, p. 7, p. 18, p. 59, p. 63, p. 96, p. 119, p. 123 and should also be corrected in these instances.			

Issue 2 End of life criteria

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 8 Sentence: 'On the basis of the results of the 35 parametric OS models considered within the ERG's exploratory analyses, the ERG does not believe that blinatumomab meets NICE's criteria for life-extending treatments given at the end of life.' P.120–121 Sentence: 'The ERG strongly disagrees with the company's proposed use of median values to determine whether NICE's end of life criteria are met. Medians represent the "middle patient" and do not take account of skewness in the distribution of patient outcomes; conversely, the only measure of central tendency which fully represents outcomes for the population as a whole is the mean' NICE has previously accepted the use of median OS values for the end of life criteria (TA396 ['Trametinib in combination with dabrafenib for treating unresectable or metastatic melanoma'] and TA366 ['Pembrolizumab for advanced	We request that the ERG acknowledge in the report that previous appraisals (eg. TA396 and TA366) have accepted that the use of median life expectancy rather than mean is appropriate to inform whether the End-of-Life criteria is met.	Previous appraisals have accepted the use of median OS, rather than mean OS. As such, it is overly critical and potentially misleading to suggest that blinatumomab does not meet NICE's criteria for life-extending treatments on the basis of median vs. mean survival.	This is not a factual inaccuracy. The ERG believes that the use of medians is inappropriate for informing whether NICE's end of life criteria are met. This is particulary important in instances whereby where treatments are potentially curative as mean and median OS estimates will diverge significantly due to the long tails of the distributions (as is the case in this appraisal). No amendment has been made.

melanoma not previously treated with ipilimumab']). As such, it is unfairly critical to reject the use of medians in this case without mention of previous appraisal decisions.			

Issue 3 Limitations associated with single-arm studies

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 26 Paragraph: 'Single-arm studies are low on the hierarchy of study quality as they are associated with potential biases. The absence of blinding leads to a risk of performance bias. The lack of randomisation leads to a risk of selection bias.' This critique should include the context that is provided in the company submission relating to the low incidence of this disease and its lack of standard of care treatments. In addition, the company submission specifies that the Committee for Medicinal Products for Humans (CHMP) noted, in its Scientific Advice provided on December 17th, 2009, that it could accept data from a single-arm trial if "good quality comparative controls would be available which would well match the patient population in the proposed	P.26: Amend to: 'Single-arm studies are low on the hierarchy of study quality as they are associated with potential biases. The absence of blinding leads to a risk of performance bias. The lack of randomisation **may** lead to a risk of selection bias **in comparative analyses**. **However, given the very low incidence of MRD+ BCP-ALL and lack of standard of care interventions, conducting large randomised clinical studies in this patient population is complex. In addition, in their Scientific Advice provided on December 17th, 2009, the Committee for Medicinal Products for Human Use (CHMP) stated that it could accept data from a single-arm trial if "good quality comparative controls would be available which would well match the**	Although the ERG provides fair and relevant criticism of single-arm study designs, it is important to acknowledge the context around why this study design was considered to be appropriate by regulatory authorities.	This is not a factual inaccuracy.

confirmatory study." On p. 59, the report acknowledges that 'the ERG considers the population characteristics of the BLAST PAS and the historical control DCAS to be representative of Ph- CR1 patients with MRD+ BCP-ALL'. As such, while it is valid to criticise single-arm study designs, it should be provided in context of the rationale for its use. In addition, as a point of accuracy it should be noted that single-arm and retrospective studies are not in themselves subject to selection bias, which only becomes a factor when conducting comparative analyses.	**patient population in the proposed confirmatory study." The use of a single-arm design, matched to a representative historical comparator study, is therefore justified.'**		

Section 2 – Factual Inaccuracies in Reporting and Confidentiality Marking

Issue 4 Accuracy of statement

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 3 Sentence: '[redacted].' This does not reflect the results presented in the company submission. This is repeated on p.57, p. 122	Amend to: '[redacted].'	This statement is incorrect and does not reflect the [redacted]	This has been amended to read "By the end of the core stud [redacted]." In all occurances of the statement.

Issue 5 Accuracy of statement

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 9 Sentence: 'Patients are considered to have clinically significant MRD, and are described as being MRD+, if their MRD level is greater than 1 x 10^4'	The correct MRD level is 1 x 10^{-4}, rather than 1 x 10^4. Amend to: 'Patients are considered to have clinically significant MRD, and are described as being MRD+, if their MRD level is greater than **1 x 10^{-4}**'	This statement is incorrect.	This has been amended.

Issue 6 ACIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 4 Sentence: 'Comparator data	This information should be marked as	This information is academically	This is now marked as ACIC.

relating to standard care chemotherapy were provided from one historical control study, Study 20120148 ([redacted]); this study was based on data obtained from existing clinical databases.' P.4 '…the BLAST primary analysis set (PAS, n=73) and the historical control direct comparison analysis set (DCAS, [redacted]).' This information should be highlighted.	academic in confidence.	sensitive.	

Issue 7 Accuracy of statement

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 27 Sentence: 'At the time of writing, only one of the studies (BLAST) had published effectiveness data in the form of an abstract at the ASH 56th Annual Meeting (2014). This statement is not correct: five publications were identified by the systematic literature review, two reporting on the BLAST study and three reporting on Study MT103-202.	We suggest that this statement be deleted or amended to reflect the published evidence base. Citations are provided below: 1. Topp MS, Kufer P, Gökbuget N, et. al. Targeted therapy with the T-cell-engaging antibody blinatumomab of chemotherapy-refractory minimal residual disease in B-lineage acute lymphoblastic leukemia patients results in high response rate and prolonged leukemia-free survival. *J Clin Oncol* 2011 Jun 20. 29(18):2493-8 2. Klinger M, Brandl C, Zugmaier G, et al. Immunopharmacologic response of patients with B-lineage acute lymphoblastic leukemia to continuous infusion of T cell-engaging CD19/CD3-bispecific BiTE antibody blinatumomab. *Blood*. 2012;119(26):6226-6233. 3. Topp MS, Gokbuget N, Zugmaier G, et al. Long-term follow-up of hematologic relapse-free survival in a phase 2 study of blinatumomab in patients with MRD in B-lineage ALL. *Blood*. 2012;120(26):5185-5187. 4. Goekbuget N, Dombret H, Bonifacio M, et al. BLAST: A Confirmatory, Single-Arm, Phase 2 Study of Blinatumomab, a Bispecific T-Cell Engager (BiTE®)	This statement is incorrect.	This statement has been amended to read "At the time of writing, BLAST was published in two abstracts" (referencing 4 and 5 of the citations provided by the company). Citations 1 to 3 (relating to the pilot study rather than BLAST) were not referenced in the CS, and have not been included.

	Antibody Construct, in Patients with Minimal Residual Disease B-Precursor Acute Lymphoblastic Leukemia (ALL). *Blood*. 2014;124(21):379-379 5. Gökbuget N, Dombret H, Bonifacio M, et al. Long-Term Outcomes after Blinatumomab Treatment: Follow-up of a Phase 2 Study in Patients (Pts) with Minimal Residual Disease (MRD) Positive B-Cell Precursor Acute Lymphoblastic Leukemia (ALL). *Blood*. 2015;126(23):680-680.		

Issue 8 ACIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 31 Sentence: 'The studies had similar baseline ages (BLAST median age=[redacted] years, MT103-202 mean age=[redacted]years).' This information should be highlighted.	This information should be marked as academic in confidence.	This information is academically sensitive.	This has not been marked as ACIC as the data has been published. Median age in BLAST has been published in the abstract by Goekbuget (2014) and these data were also available from the following sources: U.S. National Institutes of Health clinical trials registry Accessed 21.12.2017 (also accessed 24.2.2018) https://clinicaltrials.gov/ct2/show/results/NCT00560794

			https://clinicaltrials.gov/ct2/show/results/NCT01207388

Issue 9 ACIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 31 Sentences: 'The majority of participants were Ph-, with [redacted]Ph+ participants in each study. In the BLAST study, the majority of patients ([redacted]) were in first CR.' The values of '[redacted]' should be highlighted.	This information should be marked as academic in confidence.	This information is academically sensitive.	The value `[redacted]' has now been marked as ACIC as the pilot study is academically sensitive. The 65% has not been marked as ACIC as this refers to information published in Goekbuget (2014). For BLAST, n=5 PH+ is available at https://clinicaltrials.gov/ct2/show/results/NCT01207388

Issue 10 ACIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 32 Table 5: All values in this table should be highlighted.	This information should be marked as academic in confidence.	This information is academically sensitive.	Median time from prior treatment (range), months is marked as ACIC. Philadelphia chromosome disease status, for the pilot study has now been marked as ACIC Other data were available from Goekbuget (2014) and U.S.

			National Institutes of Health clinical trials registry Accessed 21.12.2017 (also accessed 24.2.2018) and so have not been highlighted.

Issue 11 ACIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 33 Sentence: 'Although PETHEMA is quite different from practice in England, only a small percentage of patients ([redacted]) received this regimen.' The value '[redacted] should be highlighted.	This information should be marked as academic in confidence.	This information is academically sensitive.	This has now been highlighted as ACIC.

Issue 12 ACIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 35 Sentence: '[redacted]'	This information should be marked as academic in confidence.	This information is academically sensitive.	This is already highlighted as ACIC. No amendment has been made.

Issue 13 ACIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 38 Sentence: 'All [redacted] of patients achieving MRD response did so[redacted].' - '[redacted]' and '[redacted]should be highlighted.	This information should be marked as academic in confidence.	This information is academically sensitive.	This information is publically available and so has not been marked as ACIC. Details are provided below: BLAST results from the clinical trials registry https://clinicaltrials.gov/ct2/show/results/NCT01207388 provide Percentage of Participants With a Minimal Residual Disease (MRD) Response Within the First Treatment Cycle as 77.9 % (95% confidence interval 69.1 to 85.1) Pilot study results from the clinical trials registry https://clinicaltrials.gov/ct2/show/results/NCT00560794 provides Percentage of Participants With an MRD Response After Each Treatment Cycle with first treatment cycle 80.0 % (56.3 to 94.3)

Issue 14 ACIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 38 Sentences: 'A total of ninety patients (▮) achieved MRD response after one or more cycles of blinatumomab treatment, with ▮of these patients responding within one cycle. There was a higher rate of response for patients in first CR ▮, than in second CR ▮ or third CR ▮; however, only ▮ were in third CR (see Table 5), hence results on this subgroup should be treated with caution.' All of the above highlighted sections should be highlighted in the report.	This information should be marked as academic in confidence.	This information is academically sensitive.	The `▮' has now been highlighted as ACIC. The remaining information is publically available and so has not been marked as ACIC. Details are provided below: N=88 available from Goekbuget 2015 The following data were available in the Table of the abstract by Goekbuget 2014 There was a higher rate of response for patients in first CR 82% (95% CI 72% to 90%), than in second CR 71% (95% CI 54% to 85%) or third CR 50% (95% CI 1% to 99%); however, only two patients were in third CR

Issue 15 ACIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 39 Sentence:	This information should	This information is	This information is publically available and so has not been marked as ACIC.

'Within the group of 74 Ph- patients who underwent HSCT prior to relapse, the 100-day mortality probability was [redacted]' The value '[redacted]' should be highlighted.	be marked as academic in confidence.	academically sensitive.	Details are provided below: BLAST results from clinical trials registry https://clinicaltrials.gov/ct2/show/results/NCT01207388?sect=Xa70156#outcome4 100-Day Mortality After Allogeneic Hematopoietic Stem Cell Transplant. From 74 particiants analyses 100 day mortality 7% (95% confidence interval 3 to 15)

Issue 16 ACIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 40 Table 11: All values in Change from baseline at end of core study, mean (SE) column should be highlighted.	This information should be marked as academic in confidence.	This information is academically sensitive.	This information is publically available and so has not been marked as ACIC. Details are provided below: BLAST results from clinical trial registry (outcome 9 "Change From Baseline in EORTC-QLQ-C30 Scales") https://clinicaltrials.gov/ct2/show/results/NCT01207388?sect=Xfa70156#outcome9

Issue 17 ACIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 40 Table 12: All values in Change from baseline at	This information should be marked as academic	This information is academically	This information is publically available and so has not been marked as ACIC. Details are provided below:

end of core study, mean (SE) column should be highlighted.	in confidence.	sensitive.	BLAST results from clinical trial registry (outcome 10 “ Change From Baseline in EuroQoL 5-Dimension (EQ-5D) Scales “) https://clinicaltrials.gov/ct2/show/results/NCT01207388?sect=Xgfa70156#outcome10

Issue 18 ACIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 46 Sentence: ‘Baseline characteristics for patients in Study 20120148 are presented in Table 17. Most of the patients in the DCAS were from [redacted] (see CS,1 Section B.2.9.3). The DCAS included [redacted] from the UK.’ All of the highlighted sections above should be highlighted in the report.	This information should be marked as academic in confidence.	This information is academically sensitive.	This has now been marked as ACIC.

Issue 19 ACIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 55 Sentence: After applying	This information should be marked as	This information is academically	This has now been marked as

sATT weights, the 18-month RFS probability with standard care chemotherapy, without censoring for HSCT, was [redacted] (CS Appendix L) and the median RFS was [redacted].' All of the highlighted sections above should be highlighted in the report.	academic in confidence.	sensitive.	ACIC.

Issue 20 ACIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 56 Sentence: 'The HSCT rate is higher in the BLAST study ([redacted]) than the historical control study ([redacted]), and the CS states that the comparison is vulnerable to HSCT being a confounding factor.' - The values of '[redacted]' and '[redacted]' should be highlighted.	This information should be marked as academic in confidence.	This information is academically sensitive.	This has now been marked as ACIC.

Issue 21 ACIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 57 Sentence: 'Comparator data relating to standard care	This information should be marked as academic in confidence.	This information is academically sensitive.	This has now been marked as ACIC.

chemotherapy were provided from one historical control study, Study 20120148 ([redacted]), that analysed data from existing clinical databases. ‘’ The highlighted section above should be highlighted in the report.			

Issue 22 ACIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 61 Sentence: ‘These subgroups of the full study populations are described as the historical comparator DCAS ([redacted]) and the BLAST PAS (n=73).’ The highlighted section above should be highlighted in the report.	This information should be marked as academic in confidence.	This information is academically sensitive.	This has now been marked as ACIC.

Issue 23 ACIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 76 [redacted] and [redacted] should	This information should be marked as	This information is academically	This has now been marked as

be highlighted as academic in confidence.	academic in confidence.	sensitive.	ACIC.

Issue 24 Accuracy of statement

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 78 Sentence: 'A further HRQoL decrement of -0.02 is applied to the general population health utility values to reflect long-term effects of exposure to radiotherapy, chemotherapy, and HSCT.'	Amend: 'A further HRQoL decrement of [redacted] is applied to the general population health utility values to reflect long-term effects of exposure to radiotherapy, chemotherapy, and HSCT.'	This statement is incorrect.	This has now been amended as suggested.

Issue 25 ACIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 79 [redacted] should be highlighted as commercial in confidence	This information should be marked as commercial in confidence.	This information is commercially sensitive.	This has now been marked as CIC.

Issue 26 ACIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 82: Sentence: 'The model uses	This information should be marked as	This information is academically	This has now been marked as

data on the cumulative 4-year probability of having undergone pre-relapse HSCT from the BLAST PAS (72.6%) and the ATT-[redacted] to inform the blinatumomab and standard care groups, respectively.' Text should be highlighted as above.	academic in confidence.	sensitive.	ACIC.

Issue 27 ACIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 82/3: As shown in the table, the company's model suggests that the mean number of HSCTs is higher in the blinatumomab group than the standard care group (mean HSCTs blinatumomab versus standard care - 0.79 versus [redacted]). **Table 35.**	This information should be marked as academic in confidence.	This information is academically sensitive.	This has now been marked as ACIC.

Treatment group	**Mean number of HSCTs per patient**				
	Pre-relapse	**Post-relapse**	**Total**		
Blinatumomab	[redacted]	[redacted]	[redacted]		
Standard care	[redacted]	[redacted]	[redacted]		
Data should be highlighted as shown.					

Issue 28 ACIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 120 Table 51 Sentence: 'Median OS, using ATT-weighted propensity score matching analyses (Section B.2.9.4), was not estimable after more than 40 months follow-up for blinatumomab thus demonstrating a [redacted] OS survival [redacted] when compared to standard care.' - 'Not estimable' should be highlighted.	This information should be marked as academic in confidence.	This information is academically sensitive.	This has now been marked as ACIC. [redacted] and [redacted] have also been hightlighted.

Issue 29 ACIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
P. 122 Sentence: 'One historical control study (Study 20120148) of standard care chemotherapy was included ([redacted]); this study that analysed data from existing clinical databases.'	This information should be marked as academic in confidence.	This information is academically sensitive.	This has now been marked as ACIC.