NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Obinutuzumab for untreated advanced follicular lymphoma [ID1020]

Committee Papers

NICE National Institute for Health and Care Excellence

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Obinutuzumab for untreated advanced follicular lymphoma [ID1020]

Contents:

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

Pre-meeting briefing

Obinutuzumab for untreated advanced follicular lymphoma

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

1

Key issues - decision problem

- **Comparators:** Company excluded 3 comparators listed in the scope from its decision problem:
  1. rituximab monotherapy
  2. rituximab-based chemotherapy **without** rituximab maintenance
  3. bendamustine monotherapy
  - Are these treatment used for previously untreated follicular lymphoma in the NHS? Is rituximab maintenance treatment routinely used after induction with rituximab-based chemotherapy?
- **Chemotherapy regimens during induction:** Only 3 chemotherapy regimens were given in combination with obinutuzumab (intervention) or rituximab (comparator) during induction phase. These were
  1. Bendamustine
  2. CHOP [cyclophosphamide, doxorubicin, vincristine and prednisolone]
  3. CVP [cyclophosphamide, vincristine and prednisolone])
  - Would other chemotherapy regimens such as MCP (mitoxantrone, chlorambucil and prednisolone) or CHVPi (cyclophosphamide, doxorubicin, etoposide, prednisolone and interferon-α) be used for induction?

2

Key issues clinical-effectiveness

- What is the appropriateness of the evidence base given that GALLIUM (the key trial):
  - Is open label and 1◦ outcome is investigator assessed progression-free survival
  - Is immature: only 7.9% of people died during the trial
  - Did not randomise the chemotherapy accompanying obinutuzumab or rituximab for induction. Instead, they were trial-site specific
  - Has a different proportion of people receiving CHOP, CVP or bendamustine compared with UK practice
  - Has younger participants than the UK patient population which affects cost effectiveness estimates
  - Is not complete: trial is ongoing

3

Key issues - cost-effectiveness (1)

- **Which progression-free survival (PFS) data?**
  - Company used investigator-assessed PFS; ERG considered this prone to bias and less reliable than independent review committee (IRC) assessed progression-free survival → this is a driver of cost effectiveness
- **Which progression-free survival probability distribution?**
  - ERG preferred a Weibull curve fitted to IRC-PFS data over an exponential curve fitted to INV-PFS used by the company
- **How long is the treatment effect?**
  - In absence of long-term data, company assumed that PFS benefit with obinutuzumab maintained until 9 years (based on rituximab in another study). ERG considered this 'speculative'
  - Considering a duration of treatment effect <5 years worsened the ICER of obinutuzumab compared with rituximab to >£30k/QALY in ERG base-case

4

Key issues - cost-effectiveness (2)

- **Estimating mortality**: To estimate mortality from states of progression-free and early progression, company pooled deaths in both arms of GALLIUM and used **same** mortality rates for both treatment arms. ERG preferred different values per treatment arm. Which is better?
- **Cost of comparator**: Should this appraisal consider low-cost biosimilars for rituximab, the comparator?
- **Utility**: ERG considers company's source of utility to be "unpublished, inconsistent with the results of the GALLIUM trial and unverifiable"

5

Obinutuzumab

Positive opinion CHMP
(Committee for Medicinal Products for Human Use, EMA)

Obinutuzumab (Gazyvaro, Roche)	
Mechanism of action	• Type II anti-CD20 antibody • targets the CD20 on non-malignant and malignant pre-B and mature B-lymphocytes • spares haematopoietic stem cells, pro-B cells, normal plasma cells or other normal tissue
Marketing authorisation	• CHMP positive opinion for: 'Obinutuzumab in combination with chemotherapy, followed by obinutuzumab maintenance therapy **in patients achieving a response**, is indicated for the treatment of patients with **previously untreated advanced** follicular lymphoma.'

6

CHMP:
http://www.ema.europa.eu/docs/en_GB/document_library/Summary_of_opinion/human/002799/WC500231836.pdf

Existing MA (subject of another NICE appraisal)
with bendamustine followed by obinutuzumab maintenance therapy for follicular lymphoma that did not respond or progressed during or up to 6 months after treatment with rituximab or a rituximab-containing regimen

FAD ID841: Obinutuzumab in combination with bendamustine followed by obinutuzumab maintenance is recommended for use within the Cancer Drugs Fund as an option for treating follicular lymphoma that did not respond or progressed during or up to 6 months after treatment with rituximab or a rituximab-containing regimen, only if the conditions in the managed access agreement for obinutuzumab are followed.

Obinutuzumab

Administration and dose

Dose	1000 mg (fixed)
Administration	Intravenous
Frequency	
Induction	With CHOP (cyclophosphamide, doxorubicin, vincristine and prednisone) or with CVP (cyclophosphamide, etoposide, doxorubicin and prednisolone): 21 day cycle 1st cycle: on day 1,8 and 15 2nd to 8th cycle: on day 1
	With bendamustine (28-day cycle) 1st cycle: on day 1,8 and 15 2nd to 6th cycle: on day 1
Maintenance in those responding to induction	Once every 2 months up to 2 years or until progression
Average course	6–8 cycle induction then up to 12 doses for responders to induction

7

Follicular lymphoma

- 2nd most common non-Hodgkin lymphoma (NHL) in Western Europe and United States
    - 35% of all NHLs
- UK incidence 3.3 per 100,000 per year
- 2,142 new diagnosis in England (2015)
- Prevalence (10 year-UK): 25.7 per 100,000
- Risk factors: use of immunosuppressive, age, sex, life style
- Male: Female ratio: 0.9
- Median age at diagnosis in UK ~65 years
- Median life expectancy 8–12 years (to 15 years- after-rituximab)
- Early progression (within 2 years) associated with increased risk of death

8

Source: Company submission, section 3.2

Follicular lymphoma grade

- Typical presenting symptom: lymph nodes enlarged at multiple sites
- Other symptoms; fatigue, weight loss, fever and night sweats
- Grading done by histological examination of surgical specimen/biopsy (based on number of centroblast*/high power field)

Grade	Description
1	≤5 blasts/high power field
2	6-15 blasts/high power field
3A	>15 blasts/high power field, centroblasts with intermingled centrocytes**
3B	>15 blasts/high power field, pure sheets of blasts

* Centroblast; an enlarged and proliferating activated B cell
** Centrocyte: the result of proliferating centroblasts

9

Source: Company's submission Table 9 page 34

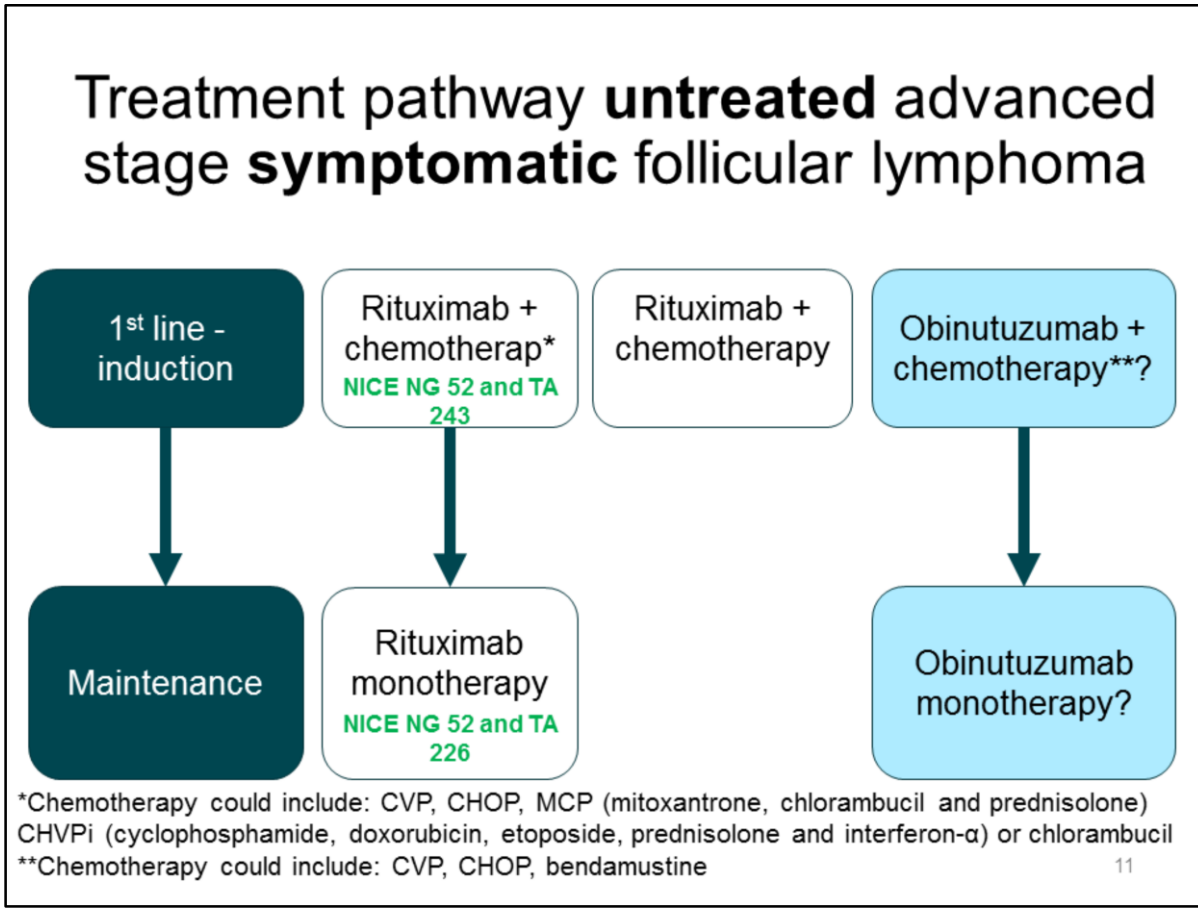
Follicular lymphoma stages

- Staging: Ann-Arbor Classification
- Stage III-IV comprise advanced disease

Stage	Area of involvement
I (I_E)	1 lymph node region or extralymphatic (IE) site
II (II_E)	≥ 2 lymph node regions or at least 1 lymph node region + 1 localised extralymphatic site (IIE) on same side of diaphragm
III (III_E, III_S)	Lymph node regions or lymphoid structures (e.g. thymus, Waldeyer's ring) on both sides of the diaphragm with optional localised extranodal site (IIIE) or spleen (IIIS)
IV	Diffuse or disseminated extralymphatic organ involvement
For all stages	
A	No symptoms
B	*Unexplained fever of >38°C, drenching night swears; or loss of >10% body weight within 6 months

10

Source: Company's submission Table 10 (page 35)



Based on Company's submission Figure 4 (page 41)

Decision problem (population and intervention)

	Scope	Decision problem	Rationale if different
Population	People with untreated advanced follicular lymphoma		-
Intervention	Obinutuzumab combined with chemotherapy, **with or without** obinutuzumab maintenance	Obinutuzumab combined with chemotherapy (CVP, CHOP or bendamustine), **followed by** obinutuzumab maintenance in patients achieving a response	Aligned with anticipated marketing authorisation. Company did not present evidence **without** obinutuzumab maintenance
ERG comments: Trial excluded people with (histological) grade 3b follicular lymphoma Trial mandates chemotherapy with CVP, CHOP or bendamustine. ERG agrees with excluding obinutuzumab + chemotherapy, **without** obinutuzumab maintenance			

grade 3b lymphomas are likely to grow faster and usually treated in the same way as diffuse large B-cell lymphoma (UK Lymphoma Association)

Decision problem - comparators

Scope	Decision problem	Rationale if different
• Rituximab monotherapy (off-label) • Rituximab-based chemotherapy, with or without rituximab maintenance • Bendamustine monotherapy (off-label but funded via the CDF)	• Rituximab in combination with chemotherapy, followed by rituximab maintenance in patients achieving a response	• Rituximab **without chemotherapy** → induction treatment only for **asymptomatic** disease • Rituximab-based chemotherapy, **without** rituximab maintenance not clinical practice • UK Systemic Anti-Cancer Therapy Dataset and market research show little use of bendamustine monotherapy

ERG comments:

1. Re maintenance: anticipated marketing authorisation includes maintenance therapy, so, rituximab-based chemotherapy without rituximab maintenance treatment can be ignored as a comparator
2. Company should have included evidence of obinutuzumab vs. rituximab mono-therapy and bendamustine monotherapy

13

Decision problem - outcomes

	Scope	Decision problem	Rationale if different
Outcome(s)	• overall survival • progression-free survival • overall response rate • adverse effects of treatment • health-related quality of life		-

ERG comments:

- Company provides all outcomes specified in the scope.
- OS data are still immature with 7.9% having died at the updated analysis cut-off date (10 September 2016) of the GALLIUM, with less than 20% of patients followed for survival for more than four years

14

Professional and clinical expert feedback

- Follicular lymphoma runs a chronic relapsing course requiring multiple episodes of treatment and culminating in resistance to therapy and/or large-cell transformation
- Median progression-free is 6 to 8 years and overall survival is 12 to 15 years
- Quality of life and time to next treatment are important considerations for patients and clinicians
- Initial treatment for advanced-stage consists of 6-8 cycles of rituximab combined with one of several different chemotherapy regimens
- For patients who achieve an anatomical complete or partial response then maintenance therapy with rituximab alone is an option (recommended in TA226)

15

Professional and clinical expert feedback

- Clinical opinion differs on **rituximab maintenance** for 3 main reasons:

1. Questionable effectiveness
   - data from the PRIMA trial indicates that the benefit of rituximab maintenance compared to no maintenance occurs during and shortly after the 2-year maintenance and consists of a delay in disease progression in only about 1 in 5 patients and delays need for further chemotherapy in only about 1 in 10 patients
   - rituximab maintenance does not prolong survival
2. Increases risk of infection
3. Increase in use of blood products
   - A large meta-analysis and a population-based study showed an increase in blood transfusion and growth factor usage in patients receiving maintenance treatment

16

Professional and clinical expert feedback

- Obinutuzumab takes longer to infuse than rituximab
- Rituximab can be given subcutaneously
- Cheaper biosimilars of rituximab now available
- In GALLIUM, the **absolute** difference in 3-year PFS between obinutuzumab & rituximab is only 4% (77.9% vs 81.9%) as assessed by an independent review committee
- Compared with rituximab, obinutuzumab was also associated with more grade ≥3 infections (20% vs 15.6%), infusion-related reactions (12.4% vs 6.75) and 2nd malignancies (4.7% vs 2.7%)
- More infusion reactions, neutropenia and infection rates from obinutuzumab could impair the quality of life of patients in remission

17

Professional and clinical expert feedback

- 1 clinical expert highlighted an 'unexpected' result from GALLIUM
  - patients who received bendamustine in combination with either rituximab or obinutuzumab had a high death rate

'*Bendamustine is not approved as frontline treatment for FL but is nevertheless widely used for this indication in combination with rituximab. This observation calls into question the use of bendamustine as a chemotherapy partner for both rituximab and obinutuzumab in this setting*'

- Clinical experts noted that in the GALLIUM trial the dose of obinutuzumab was higher than that of rituximab.
  - Rituximab at 375 mg/m^2 (n.b. 1.7 m^2 is an approximate body surface area) at each administration whereas obinutuzumab was given at a flat dose of 1000 mg.
  - Obinutuzumab was given on day 1, 8 and 15 of cycle 1 as well as day 1 of each subsequent induction cycle, whereas rituximab was given only once within each induction cycle.

18

Clinical evidence: 1 key trial GALLIUM

- Ongoing, phase III, multicentre, open-label, randomised controlled trial
- Asked question:
  - in people with follicular lymphoma (grade 1 to 3a), does obinutuzumab-chemotherapy induction followed by obinutuzumab maintenance delay progression of disease compared with rituximab-chemotherapy followed by rituximab maintenance treatment?
- Each site chose 1 of 3 chemo- regimens (CHOP, CVP, or bendamustine), and all patients at a given site received the same chemotherapy regimen in combination with obinutuzumab or rituximab for induction
- GALLIUM used by company to model:
  - time to progression
  - time on treatment
  - post progression survival for people who progress early (post progression survival for late progression see next slide)
  - n.b. time to death modelled

19

Other trial evidence GAUDI and PRIMA trials

GAUDI: randomised open label phase I b study

- a sub-study of GAUDI (Grigg et al., 2017) compared safety and efficacy of 2 induction regimen in patients with previously untreated follicular lymphoma:
  - obinutuzumab-CHOP (cyclophosphamide, doxorubicin, vincristine and prednisone) with obinutuzumab-bendamustine
  - both followed by obinutuzumab maintenance
  - **includes obinutuzumab**; **not used in modelling**

PRIMA randomised phase III study compared rituximab maintenance therapy with observation only:

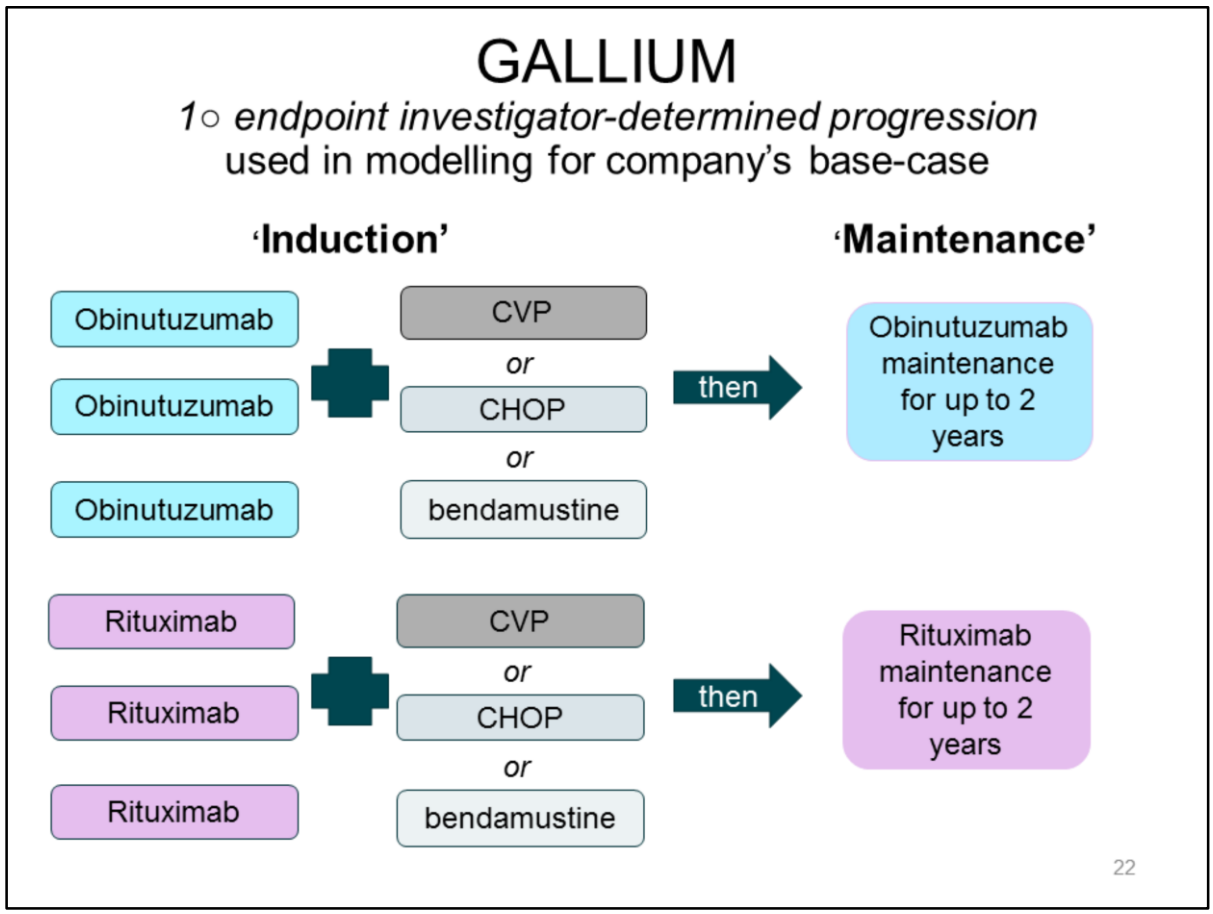
- in people with in previously untreated follicular lymphoma, following induction with rituximab+ chemotherapy
- Follow-up data up to 9.75 years
- **Does not include obinutuzumab, but used in modelling**
  - To populate time from late progression to death for both obinutuzumab and rituximab

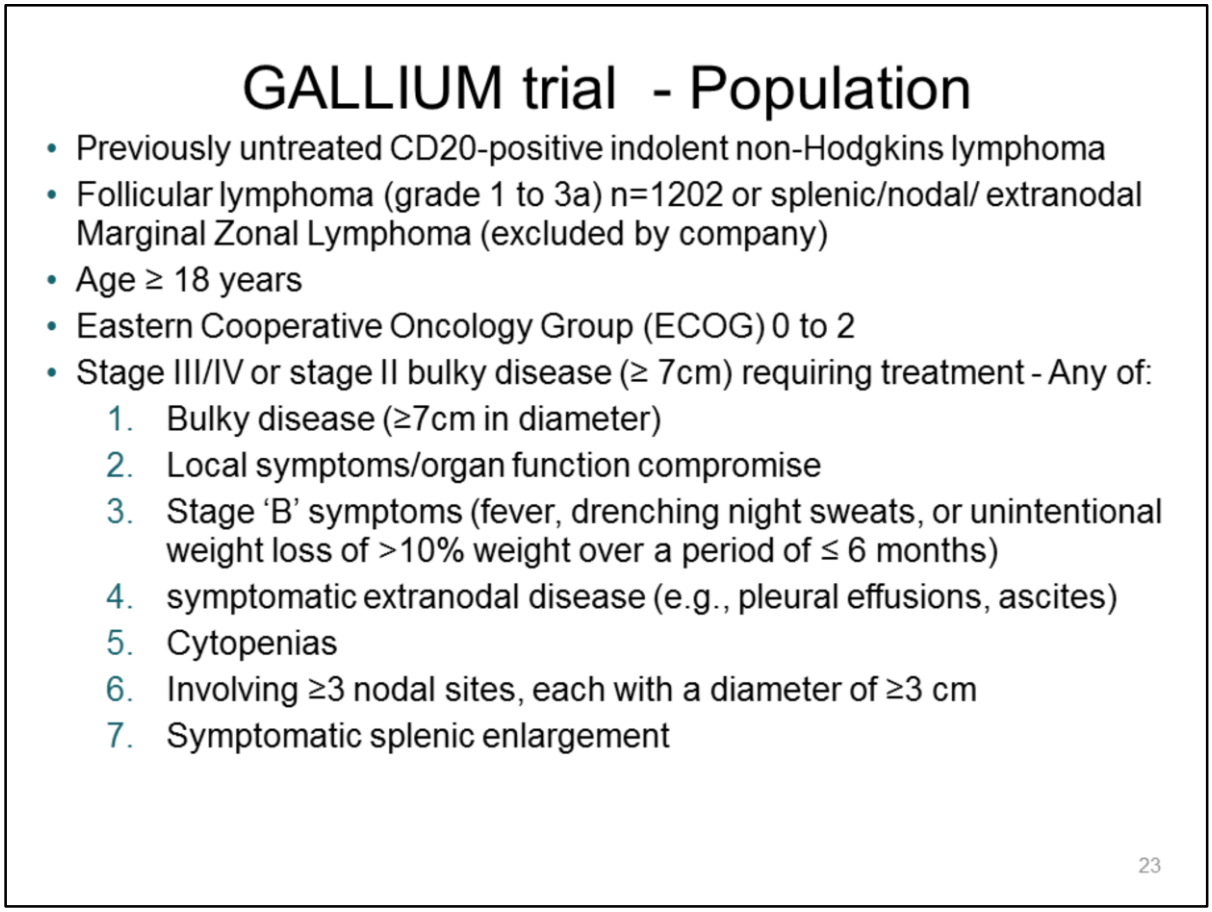
20

GALLIUM trial: overview

- Ongoing, phase III, open-label, randomised controlled trial
- 1202 patients with follicular lymphoma + 199 patient with marginal zonal lymphoma
  - Note: only patients with follicular lymphoma included in the evidence presented by the company
- Multicentre: 177 trial centres in 18 countries
- 293 patients (21%) of patients were from the UK
- Each site chose 1 of 3 chemo- regimens (CHOP, CVP, or bendamustine) as standard of care for follicular lymphoma
  - all patients with follicular lymphoma at that site received the same chemotherapy regimen for the duration of the study

21


GALLIUM
1○ endpoint investigator-determined progression
used in modelling for company's base-case
'Induction'
'Maintenance'
Obinutuzumab
Obinutuzumab
Obinutuzumab
CVP
or
CHOP
or
bendamustine
then
Obinutuzumab maintenance for up to 2 years
Rituximab
Rituximab
Rituximab
CVP
or
CHOP
or
bendamustine
then
Rituximab maintenance for up to 2 years
22



Source: Tables 17 and 22 of the CS[1]

ECOG = Eastern Cooperative Oncology Group; FL = follicular lymphoma; iNHL = indolent non-Hodgkin lymphoma; IRC = independent review committee; MZL = marginal zone lymphoma; Obin-chemo = obinutuzumab with chemotherapy as induction, R-chemo = rituximab with chemotherapy as induction,

GALLIUM intervention and comparator

Obinutuzumab	Rituximab
8-10 doses of obinutuzumab at 1000 mg IV • O-CHOP: O on Days 1, 8, and 15 of Cycle 1 and on Day 1 of Cycles 2–8 (21-day cycles). • O-CVP: on Days 1, 8, and 15 of Cycle 1 and on Day 1 of Cycles 2–8 (21-day cycles). • O-bendamustine: O on Days 1, 8, and 15 of Cycle 1 and on Day 1 of Cycles 2–6 (28-day cycles).	**6-8** doses of rituximab at 375 mg/m^2 IV • R-CHOP: R on Day 1 of cycles 1–8 (21-day cycles). • R-CVP: R on Day 1 of Cycles 1–8 (21-day cycles). • R-bendamustine: R on Day 1 of Cycles 1–6 (28-day cycles).
Patients who achieved a complete response or partial response at the end of induction had maintenance every 2 months until disease progression, or for 2 years (max).	
CHOP on Day 1, with prednisone/prednisolone/methylprednisolone also on Days 2–5 of Cycles 1–6 CVP was administered on Day 1, with prednisone/prednisolone/methylprednisolone also administered on Days 2–5 of Cycles 1–8 Bendamustine on Days 1 and 2 of Cycles 1–6, with prednisone/prednisolone/methylprednisolone on Day 1 of Cycle 1	

24

GALLIUM Outcomes

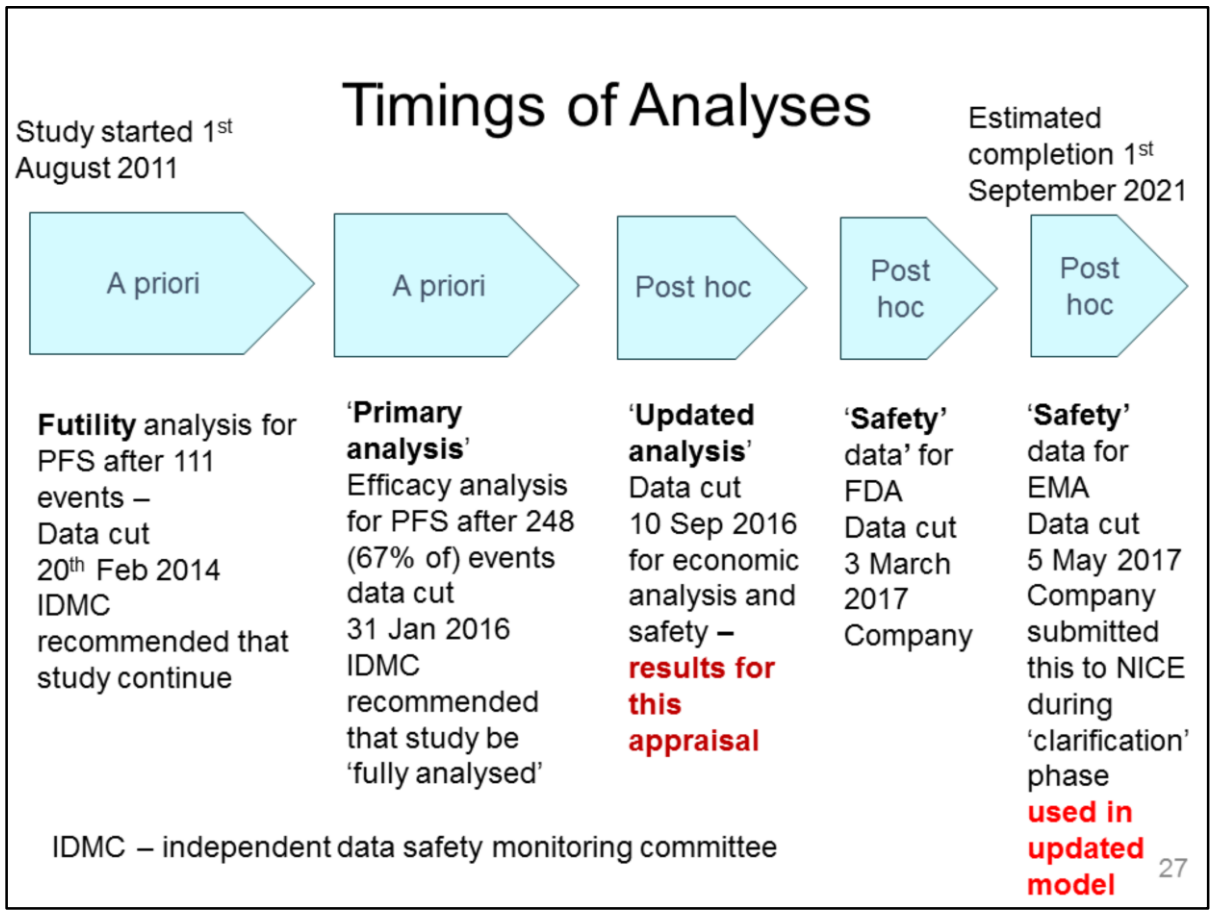
1◦ outcome	Progression Free Survival investigator assessed
2◦ outcomes	Progression-free survival (independent review) Overall survival Best overall response Disease-free survival Event-free survival Duration of response Minimal residual disease End-of-maintenance response Adverse events
Quality of life	FACT-Lym EQ-5D
Induction period only	Complete response End-of-treatment overall response

25

GALLIUM: statistical plan

NCT01332968	
False positive rate	2-sided stratified log rank test at an overall 5% significance level
Power	80% power to detect a hazard ratio (HR) for O-chemo versus R-chemo of 0.74, corresponding to an improvement in 3-year PFS from 70.7% to 77.4% or in median PFS from 6 years to 8.1 years (35%)
1◦ outcome	investigator-assessed progression free survival (INV-PFS) defined as the time from the day of randomisation until the first documented day of disease progression, symptomatic deterioration, disease transformation, or death from any cause, whichever occurred first.
Subgroup analyses	FLIPI (Follicular Lymphoma Interntional Prognostic Index), chemotherapy, geographic region

26


Timings of Analyses
Study started 1st August 2011
Estimated completion 1st September 2021
A priori
A priori
Post hoc
Post hoc
Post hoc
Futility analysis for PFS after 111 events – Data cut 20th Feb 2014 IDMC recommended that study continue
'Primary analysis' Efficacy analysis for PFS after 248 (67% of) events data cut 31 Jan 2016 IDMC recommended that study be 'fully analysed'
'Updated analysis' Data cut 10 Sep 2016 for economic analysis and safety – results for this appraisal
'Safety' data' for FDA Data cut 3 March 2017 Company
'Safety' data for EMA Data cut 5 May 2017 Company submitted this to NICE during 'clarification' phase used in updated model
IDMC – independent data safety monitoring committee
27

GALLIUM: baseline characteristics

Domain	O-chemo (n = 601)	R-chemo (n = 601)
Mean age, years (SD)	58.2 (11.5)	57.7 (12.2)
Male, n (%)	283 (47.1)	280 (46.6)
Mean body surface area, m2 (SD)	1.86 (0.2)	1.84 (0.2)
Mean BMI, kg/m2 (SD)	26.8 (5.3)	26.4 (5.9)
Race, n (%)		
Caucasian	487 (81.0)	481 (80.0)
Black or African American	3 (0.5)	1 (0.2)
Asian	100 (16.6)	98 (16.3)
Other	10 (1.7)	17 (2.8)
Geographic region, n (%)		
Eastern Europe	78 (13.0)	79 (13.1)
Western Europe	294 (48.9)	286 (47.6)
North America	75 (12.5)	77 (12.8)
Asia	92 (15.3)	93 (15.5)
Other	62 (10.3)	66 (11.0)
Chemotherapy regimen, n (%)		
Bendamustine	345 (57.4)	341 (56.7)
CHOP	195 (32.4)	203 (33.8)
CVP	61 (10.1)	57 (9.5)

Source: Table 43 ERG report, page 40-41

GALLIUM: baseline characteristics

Domain	O-chemo (n = 601)	R-chemo (n = 601)
ECOG PS, n (%)	n=600	n=599
0–1	585 (97.5)	576 (96.2)
2	15 (2.5)	23 (3.8)
Ann Arbor Stage, n (%)	n=598	n=597
I	10 (1.7)	8 (1.3)
II	41 (6.9)	44 (7.4)
III	208 (34.8)	209 (35.0)
IV	339 (56.7)	336 (56.3)
Bone marrow involvement, n/patients (%)	318/592 (53.7)	295/598 (49.3)
Extranodal involvement, (%)	392/601 (65.2)	396/601 (65.9)
Bulky disease at baseline (6 cm threshold) (%)	255/600 (42.5)	271/600 (45.2)
Mean time from diagnosis to randomisation, months (range)	6.25 (0.1–121.6)	7.28 (0.0–168.1)

29

Source: Table 43 ERG report, page 40-41

GALLIUM: baseline characteristics

Domain	O-chemo (n = 601)	R-chemo (n = 601)
FLIPI 1 (Follicular Lymphoma International Predictive Index 1) based on Age, Haemoglobin, Serum LDH, Ann Abror Stage and No of nodal site involvement		
No. of adverse factors categories 1, n (%)	n=601	n=601
Low (0,1)	128 (21.3)	125 (20.8)
Intermediate (2)	224 (37.3)	223 (37.1)
High (≥3)	249 (41.4)	253 (42.1)
FLIPI 2 (based on Age, Haemoglobin, Serum β-2 macroglobulin, Bone marrow involvement, Diameter of largest lymph node)		
No . of adverse factors categories 2, n (%)	n=579	n=586
Low (0,1)	51 (8.8)	55 (9.4)
Intermediate (2)	296 (51.1)	290 (49.5)
High (≥3)	232 (40.1)	241 (41.1)

Source: Table 43 ERG report, page 40-41

GALLIUM: baseline characteristics by chemotherapy

n (%)	Bendamustine n=686	CHOP n=399	CVP n=117
Median age, years (range)	59 (23–88)	58 (31–85)	59 (32–85)
Age ≥80 years n (%)	**23 (3.4)**	3 (0.8)	**4 (3.4)**
Male	332 (48.4)	177 (44.4)	54 (46.2)
Charlson Comorbidity Index score ≥1	163 (23.8)	69 (17.3)	22 (18.8)
ECOG PS 2	24 (3.5)	8 (2.0)	6 (5.1)
FLIPI high risk (≥3)	274 (39.9)	**187 (46.9)**	41 (35.0)
Bulky disease (≥7cm)	274 (39.9)	**206 (51.6)**	46 (39.3)

31

.

Source: Table 4.4 ERG report (page 42)

ERG comments - baseline characteristics

- Baseline characteristics appear balanced between groups
- People with African or Caribbean family origin underrepresented in the trial (N=4)
- Patients younger (median age 58 years) compared with UK patient population (median age at diagnosis 65 years)
- Proportion of patients receiving bendamustine, CHOP and CVP as part of induction (57%, 33% and 10%) may not reflect UK practice (36%, 29% and 22%)
  - Company did not randomise people to chemotherapy, so there may be confounding differences in baseline patient characteristics between the chemotherapy subgroups
  - High risk patients are more likely to receive CHOP whereas older patients more likely to receive bendamustine and CVP

32

CONFIDENTIAL

GALLIUM: Progression free survival results

Obinutuzumab improves progression free survival

	Updated analysis (cut-off September 2016)	
	Obin-chemo n=601	R-Chemo n=601
Progression-free survival (investigator-assessed)		
N (%)	120 (20.0)	161 (26.8)
Median PFS, (95% CI)	Not estimated	Not estimated
HR (stratified), 95% CI	0.68 (0.54 to 0.87)	
Progression-free survival (independently reviewed-assessed)		
N (%)	108 (18.0)	141 (23.5)
Median PFS, (95% CI), m	Not estimated	Not estimated
HR (stratified), 95% CI	0.72 (0.56 to 0.93)	

ERG comments:

- Because of open label design progression free survival results by independent review committee will be less prone to bias

33

PFS was compared using a two-sided log-rank test stratified by chemotherapy regimen (CHOP, CVP, or bendamustine), FL international prognostic index (FLIPI) risk group (low, intermediate, or high)

CONFIDENTIAL

GALLIUM: Overall survival results

Immature data

	Updated analysis (cut-off September 2016)	
	Obin-chemo n=601	R-Chemo n=601
Overall survival		
Patients w/ event, n (%)	43 (7.2%)	52 (8.7%)
Median OS, months	Not estimated	Not estimated
HR (stratified), 95% CI	0.82 (0.54 to 1.22)	

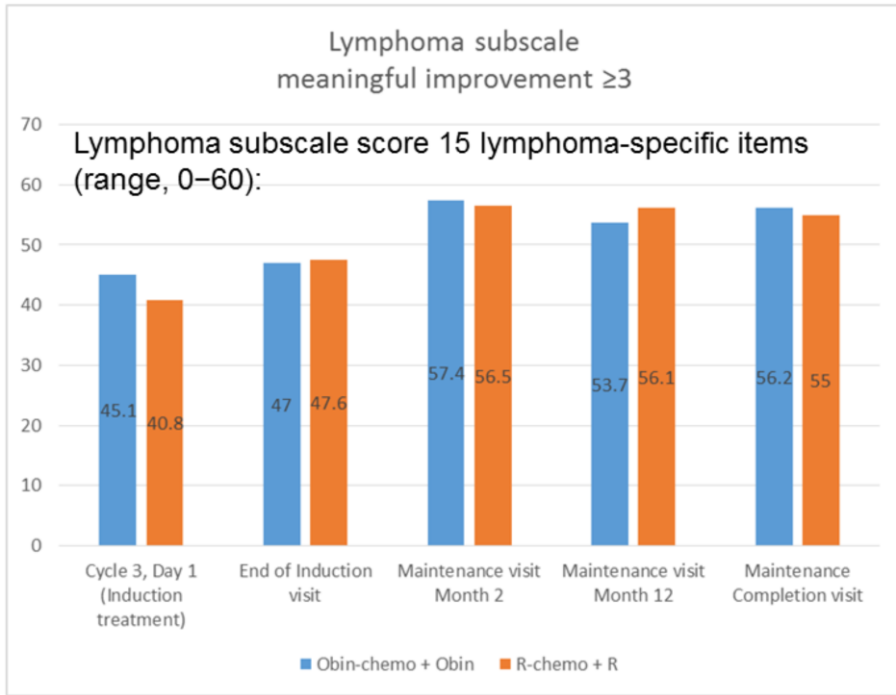
34

GALLIUM – Quality of life

- Health-related quality of life (HRQoL) data collected using 2 self-administered questionnaires:
  - Functional Assessment of Cancer Therapy for Lymphoma (FACT-Lym) and
  - EuroQol EQ-5D-3L
- The questionnaires were administered at
  - Baseline
  - Completion of induction
  - Completion of maintenance
  - Follow-up month 36
- No notable differences between the treatment arms in any of the FACT-Lym questionnaire subscales or EQ-5D-3L scales over time during the induction and maintenance treatment periods, and follow-up

35

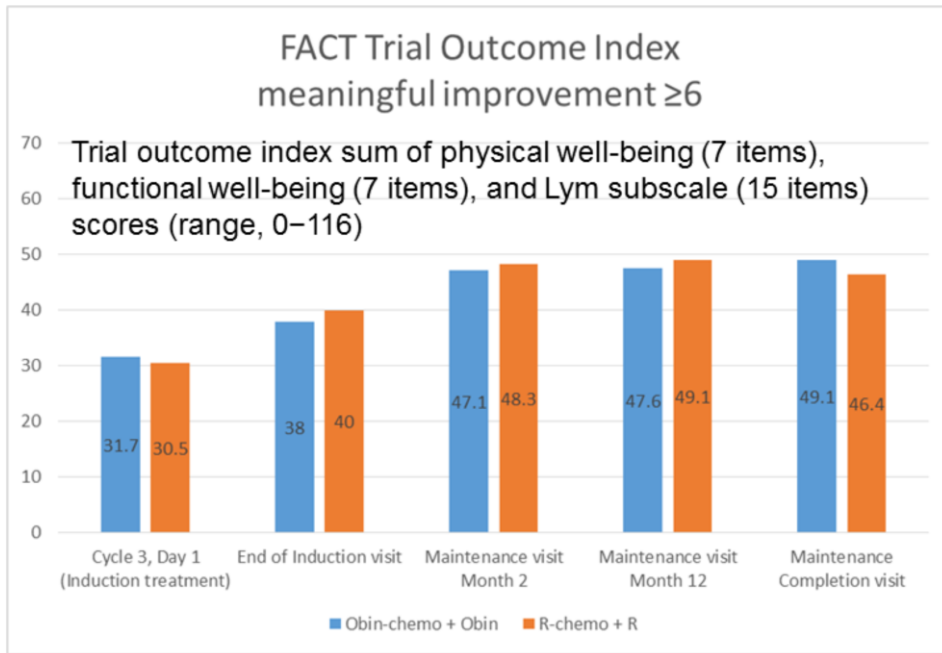
GALLIUM trial result (FACT-Lym)
% with meaningful improvement



36

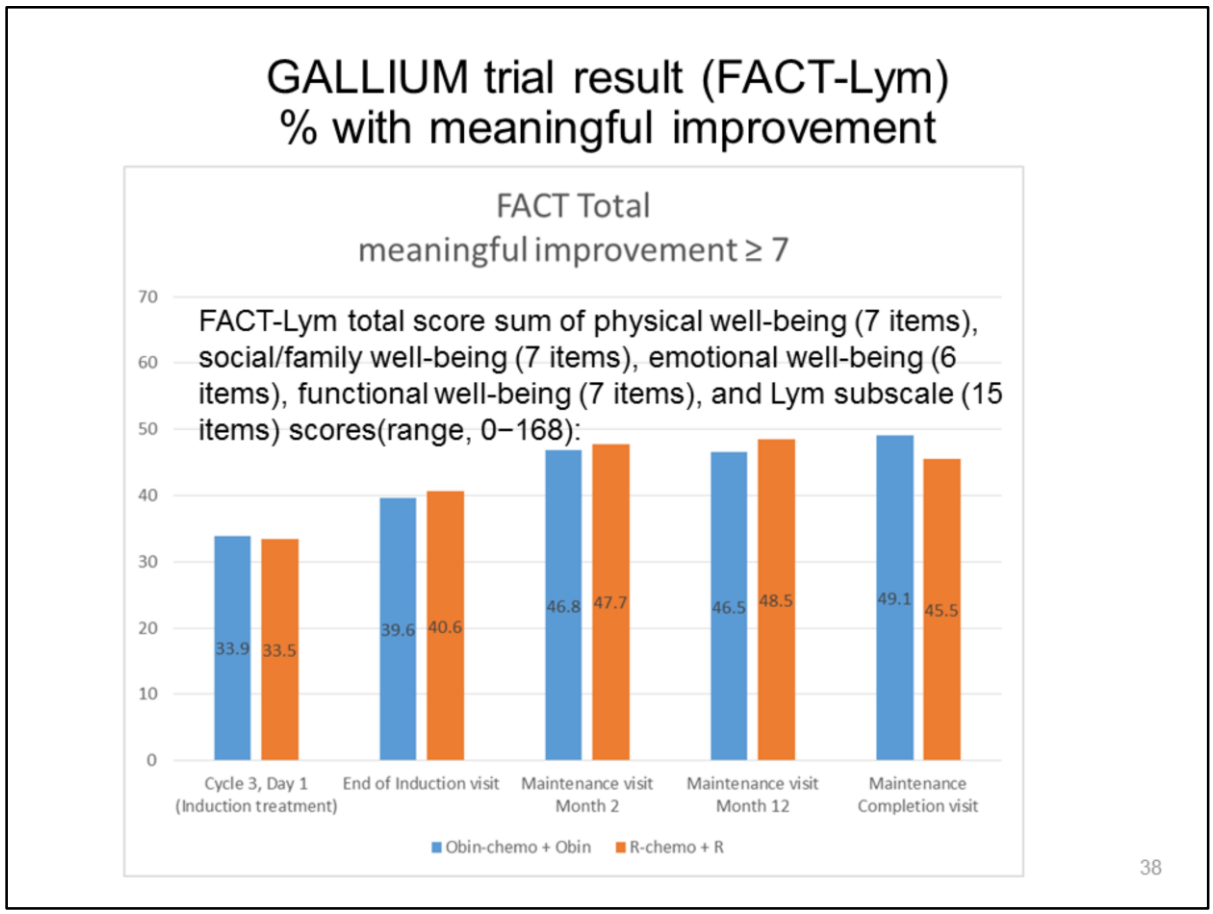
Based on Table 40 Company's submission (page 93)

GALLIUM trial result (FACT-Lym)
% with meaningful improvement



37

Based on Table 40 Company's submission (page 93)



Based on Table 40 Company's submission (page 93)

GALLIUM - EQ-5D

	G-chemo+G		R-chemo+R		Difference	
State	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	P-value
Induction off treatment	0.765	0.032	0.779	0.031	-0.015	0.72
Induction on treatment	0.823	0.015	0.824	0.015	-0.002	0.84
Maintenance & follow-up off treatment	0.826	0.015	0.810	0.015	0.017	0.13
Maintenance & follow-up on treatment	0.834	0.015	0.828	0.014	0.006	0.54
Early progression <= 2yrs	0.767	0.026	0.782	0.022	-0.015	0.62
Late progression > 2yrs	0.820	0.033	0.810	0.030	0.010	0.80

39

Source: Clarification response, Table 8 (page 26)

GALLIUM subgroups
Investigator assessed progression free survival

Subgroup	N	HR	95%CI
ITT population (patients with FL)	1202	0.66	0.51 to 0.85
FLIPI (interaction p value 0.14) No significant interaction			
Low	253	1.17	0.63 to 2.27
Intermediate	447	0.59	0.37 to 0.92
High	502	0.58	0.41 to 0.84
Chemotherapy regimen (interaction p value 0.67) No significant interaction			
CHOP	398	0.77	0.50 to 1.20
CVP	118	0.63	0.32 to 1.21
Bendamustine	686	0.61	0.43 to 0.86

40

Based on Figure 16, Company's submission (page 98)

GALLIUM subgroups

Investigator assessed progression free survival

Subgroup	N	HR	95%CI
ITT population (patients with FL)	1202	0.66	0.51 to 0.85
Geographic regions (interaction p value 0.68) No significant interaction			
Asia	185	0.46	0.22 to 0.95
Eastern Europe	157	0.71	0.36 to 1.37
North America	152	0.77	0.39 to 1.50
Other	128	0.40	0.14 to 1.12
Western Europe	580	0.73	0.51 to 1.04

41

Based on Figure 16, Company's submission (page 98)

GALLIUM subgroups

Investigator assessed progression free survival

Subgroup	N	HR	95%CI
ITT population (patients with FL)	1202	0.66	0.51 to 0.85
Sex (interaction p value **0.056**) No significant interaction			
Male	563	0.82	0.59 to 1.15
Female	639	0.49	0.33 to 0.74
Race (interaction p value 0.35) No significant interaction			
Asian	198	0.46	0.23 to 0.93
White	968	0.72	0.54 to 0.95
Other	36	0.30	0.04 to 2.52
Ann Arbor Stage (interaction p value 0.67) No significant interaction			
I	18	0.76	0.11 to 5.45
II	85	1.16	0.39 to 3.48
III	417	0.70	0.44 to 1.11
IV	675	0.59	0.43 to 0.82

Based on Figure 18, Company's submission (page 100)

ERG comments - subgroups

- Differences in progression free survival according to gender (obinutuzumab more effective in female) statistical interaction p = 0.056.
- Subgroup analyses on the basis of concomitant chemotherapy during induction indicates best results with bendamustine
  - In UK, according company's market research, most widely used chemotherapy in combination with rituximab for first-line treatment of follicular lymphoma is CVP
- The different efficacy in the subgroups based on chemotherapy may reflect
  - difference in the efficacy of chemotherapy regimen or
  - difference in patient selection

43

Safety results

- Safety data from patient with follicular lymphoma (FL) received from the primary analysis of the GALLIUM (clinical cut-off 31 January 2016).
- 1192 patients with FL received any drug during the induction phase (597 patients in the R-chemo arm, and 595 patients in the obin-chemo arm), and are included in the FL safety population
- The European Medicines Agency requested further safety analyses, which resulted in a 'Revised Safety Analysis', which is an analysis conducted on the safety data derived from a 5 May 2017 snapshot
- The company submitted revised safety data at clarification stage

44

CONFIDENTIAL

Safety Results

Date	29-Apr-16		05-May-17	
	Obin-chemo	R-chemo	Obin-chemo	R-chemo
	n = 595	n = 597	n = 595	n = 597
No. of patients with at least 1 AE (%):				
AE (all grades)	99.5	98.0	99.8 (+0.3%)	99.2 (+1.2%)
Grade 3-5 AE	74.6	67.8	76.6 (+2.0%)	70.0 (+2.2%)
Fatal AE	4.0	3.4	4.0 (+0%)	3.4 (+0%)
Serious AE	46.1	39.9	46.6 (+0.5%)	40.0 (+0.1%)
AE leading to withdrawal from any treatment	16.3	14.2	16.0 (+0.3%)	14.4 (+0.2%)

ERG comments:

- Highlighted higher rate of serious and higher grade events with obinutuzumab

45

Source: ERG report Table 4.2 (page 53)

CONFIDENTIAL

Adverse events May 2017

	Obin-chemo		R-chemo	
AEs of particular interest, (%):	All Grades	Grade ≥3	All Grades	Grade ≥3
Infusion-related reaction (IRR)	70.6	12.3	60.5	7.4
Neutropenia	52.3	47.7	47.1	41.4
Infection	80.2	20.7	72.9	16.4
Tumour lysis syndrome	1.0	1.0	0.5	0.5
Thrombocytopenia	12.4	6.1	8.0	2.8
Acute thrombocytopenia	1.2	0.7	0	0
Hemorrhagic events	11.4	1.0	11.4	1.2
GI perforation	0.8	0.5	0.5	0
Cardiac events (incl. IRRs)	14.3	4.0	10.1	2.8
(excl. IRRs)	10.6	3.4	8.5	2.5
Second malignancy (system organ class)[a]	11.1	5.0	7.5	2.8
StandardizedMedical Dictionary for Regulatory Activities query	7.6	4.7	5.2	2.7

46

SMQ = standardized MedDRA query; SOC = system organ class; TLS = tumour lysis syndrome.
Source: ERG report Table 4.2 (page 53)

Results from GAUDI trial

- Obinutuzumab-CHOP → obinutuzumab (n=41) compared with obinutuzumab-bendamustine → obinutuzumab (n=40)

Outcome	O-benda (n = 41)	O-CHOP (n = 40)	Total (n = 81)
Efficacy			
ORR (%) (95% CI)	93 (80.1 to 98.5)	95 (83.1 to 99.4)	94 (86.2 to 98.0)
CR at end of induction, (%) (95% CI)	37 (22.1 to 53.1)	35 (20.6 to 51.7)	36 (25.4 to 47.2)
CR at 30 months, (%) (95% CI)	63 (46.0 to 78.2)	58 (40.8 to 74.5)	61 (NA to NA)
PFS at 36 months, % (95% CI)	90 (0.80 to 0.99)	84 (0.72 to 0.96)	87 (0.79 to 0.94)
Progression / death (n)	6	11	17
Deaths due to PD (n)	1	2	3
Safety			
Induction Grade 3 / 4 AE n (%)	21 (51)	31 (78)	52 (64)
Maintenance Grade 3 -5 AE n (%)	NR	NR	27 of 72 (37.5)

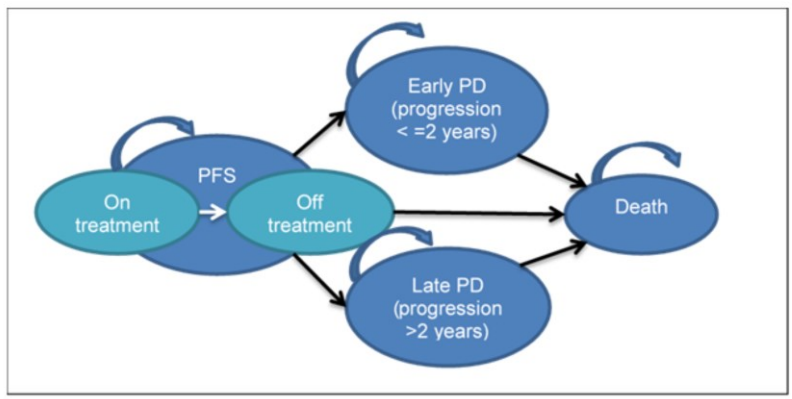
Source: Table 44 of Company's submission (page 104)

Cost effectiveness

48

Company's model

- State transition Markov model
- 4 health states,1 month cycles, 40 year time horizon (amended to 50 after clarification)
- NHS/PSS perspective; Costs and benefits discounted at 3.5%



ERG's comments:

- "structure seems appropriate for the decision problem"

49

Source: Figure 22, Company's submission (page 135)

How the model works

- Patients begin in progression free survival (on treatment) state
- Patients responding to induction receive maintenance treatment
  - Treatment continued until progression or for a maximum of 2 years
- Time to treatment discontinuation based on observation from GALLIUM
- After completion or discontinuation of treatment in the PFS, patient remain in PFS (off treatment) until progression
  - Patients can either remain in PFS or exit due to disease progression or death
- Disease progression:
  - Early (with in 2 years) and late progression disease states (EPD and LPD) were differentiated to account for worse outcome in people with early progression
  - Once patients enter any progressive-disease state, they remain in the corresponding PD state until death

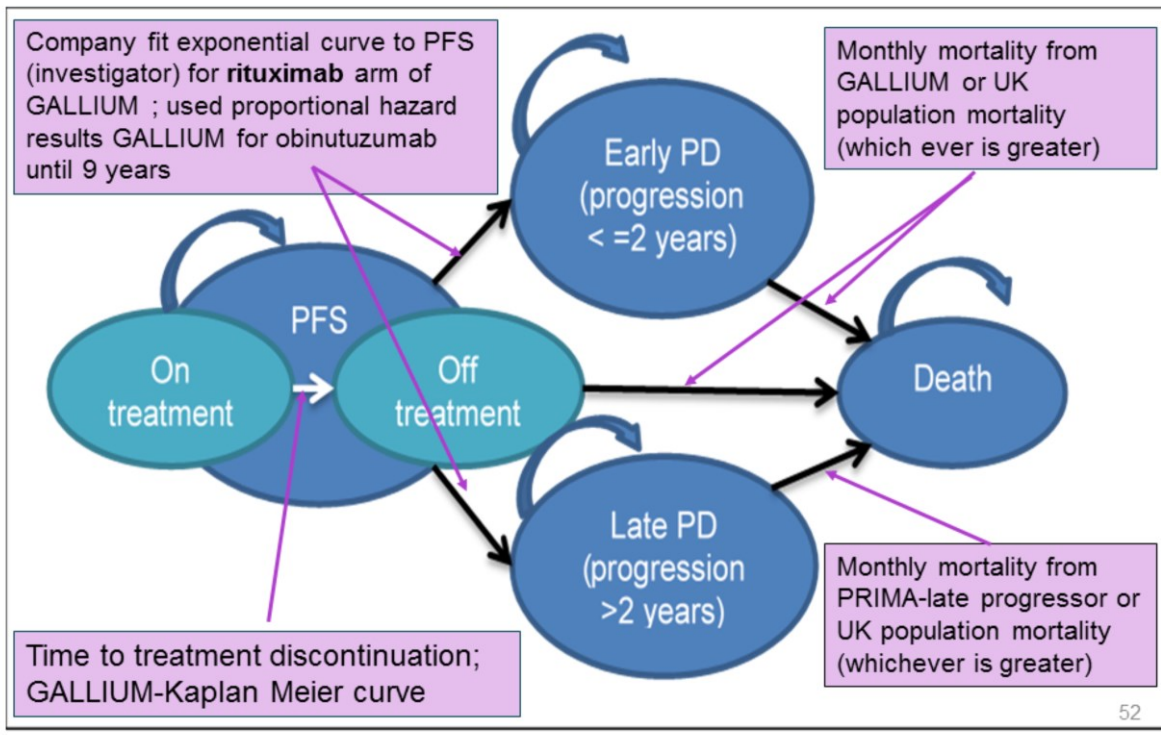
50

Summary of transitions in the model

Transition	Transition probability	
Progression free survival to early progressed disease (≤2 years) and late progressed disease (>2 years)	• Time dependent • Calculated from the probability of remaining in progression free survival and probability of death in progression free survival health state • Probability of remaining in PFS modelled with parametric model (base case Weibull) and proportional hazards	
Progression free survival to death	Based on trial mortality from GALLIUM	Greater of trial mortality or general UK population background mortality (Trial mortality applied up until age specific general UK population mortality becomes higher)
Early progressed disease (≤2 years) to death	Based on mortality from GALLIUM	
Late progressed disease (>2 years) to death	Based on mortality from PRIMA (late progressor)	

Source: Company's submission, table 62, page 140

Company source of parameters in model



How company modelled chemotherapies

- Company assumed no difference in effectiveness between the types of chemotherapy regimen

Proportion of patients taking each chemotherapy	GALLIUM (whole population) used in company's base-case	GALLIUM (UK population) ERG's preferred assumption	UK population questionnaire based UK sample (N=157, from 45 clinicians)
Bendamustine	57%	68%	29%
CHOP	33%	1%	13%
CVP	10%	31%	36%

ERG comments:
- GALLIUM not powered to detect differences between chemotherapy
- Uncertainty over proportional use of 3 chemotherapies
- ERG prefers GALLIUM UK proportions for its base-case

53

Source: ERG report page 67.

Population characteristics used in company's model

- Baseline characteristic were based on the GALLIUM Trial

Patient characteristic	Baseline value (mean)
Age (years)	57.9
Body weight (kg)	75.7
Height (cm)	168.3
Calculated Body Surface Area [BSA] (m^2)	1.86

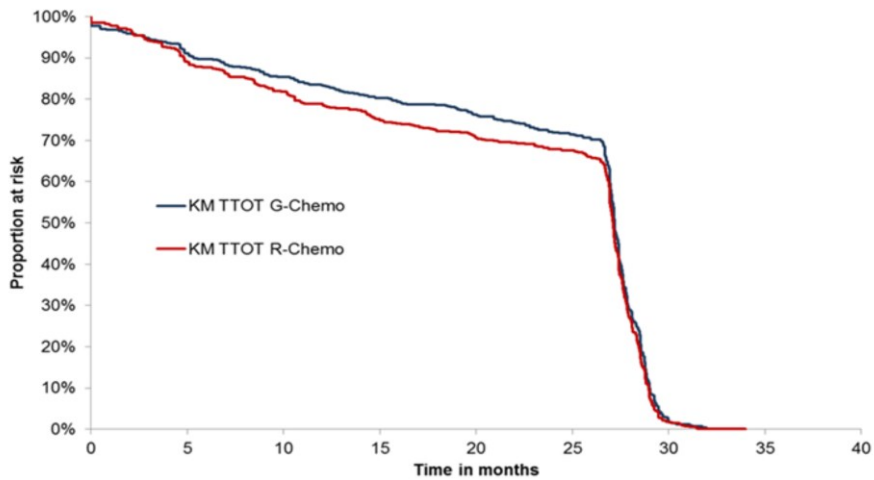
ERG's comments:

- Median age in GALLIUM 59 years; Median age at diagnosis Haematological Malignancy Research Network (HMRN) is 65 years
  - Company: HMRN relates to all follicular leukaemia patients irrespective of treatment; could include patients with less advanced disease
  - ERG: HMRN shows median age of 63.7 years for people treated with chemotherapy - **"A higher baseline age should have been used"**
- BSA: Unlikely that dosing in NHS would differ from GALLIUM – but parameter not properly coded for use in probabilistic sensitivity analysis

54

Clinical data in model: stopping treatment

- GALLIUM Kaplan-Meier curves for time-to-treatment-discontinuation
- Extrapolating not needed as all patients completed or stopped treatment
- Maintenance treatment for a maximum of 2 years



55

Source: Figure 6, Clarification response (page 32)

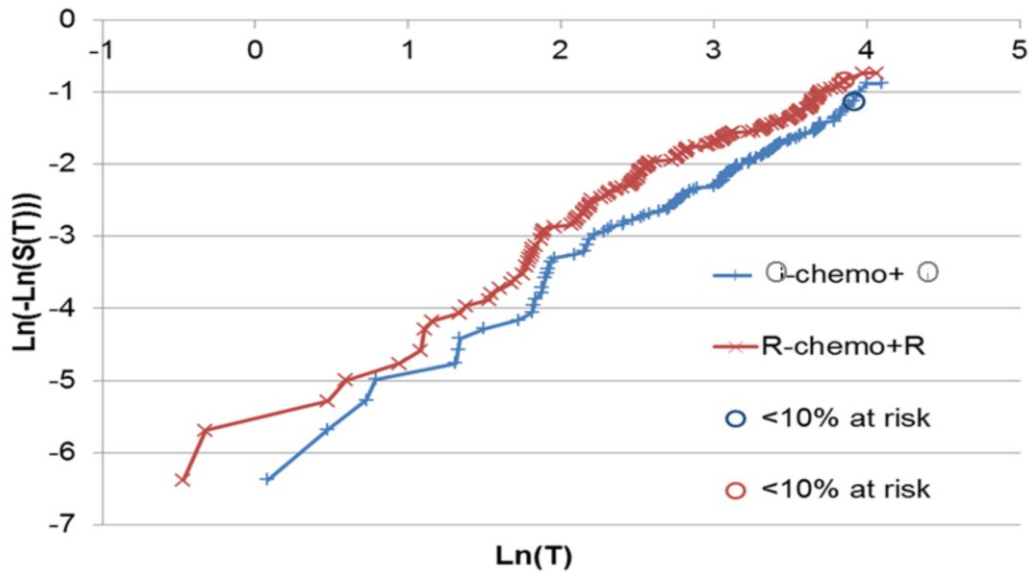
Modelled progression free survival

- Company modelled progression-free survival using investigator assessed patient level data from GALLIUM for the rituximab arm and then extrapolated with a parametric curve and then applying a constant hazard for obintuzumab's treatment effect
  - Company considered proportional hazard assumption holds by visually inspecting log-cumulative hazards plot and cumulative hazard plot (next slide)
  - Company selected exponential curve by comparing the tail of the parametric fits of the rituximab arm with long-term data from:
    - PRIMA (Phase III, RCT of rituximab maintenance in patients responding to rituximab plus chemotherapy (CHOP or CVP but no bendamustine) induction, FU data up to 9.75 years)
    - US LymphoCare registry patients receiving R-CHOP, R-CVP or R with a fludarabine-based regimen (no R-bendamustine). Median follow-up was 7.4 years
    - UK advisory board feedback that relapse rate 30-40% at 10 years

56

Proportional hazards assumption

- Log-cumulative hazard plot for PFS in GALLIUM

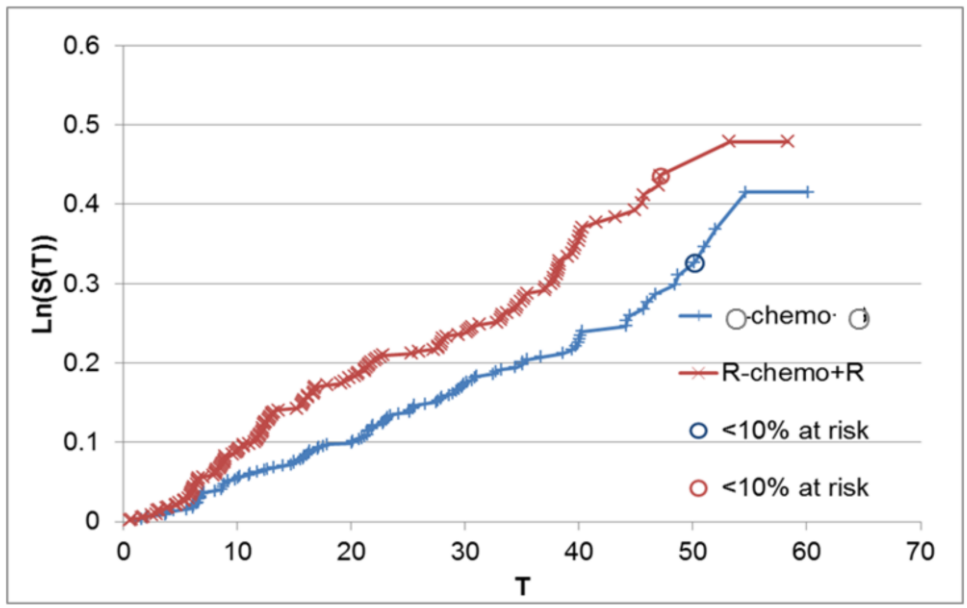


57

Source: Company's submission Figure 23 (page 141)

Proportional hazards assumption

- Cumulative hazard plot GALLIUM PFS INV



58

Source: Company's submission Figure 24 (page 142)

PFS (INV) extrapolations for rituximab arm Goodness of fit

Exponential distribution selected despite being ranked below other distributions because of external validity of long-term predictions (tail)

Distribution	AIC	Ranking	BIC	Ranking
Exponential	**1785.9**	**5**	**1796.1**	**3**
Weibull	1782.2	4	1797.5	5
Log-logistic	1779.9	3	1795.1	2
Lognormal	1774.5	1	1789.7	1
Generalised Gamma	1776.4	2	1796.8	4
Gompertz	1785.9	6	1801.2	6

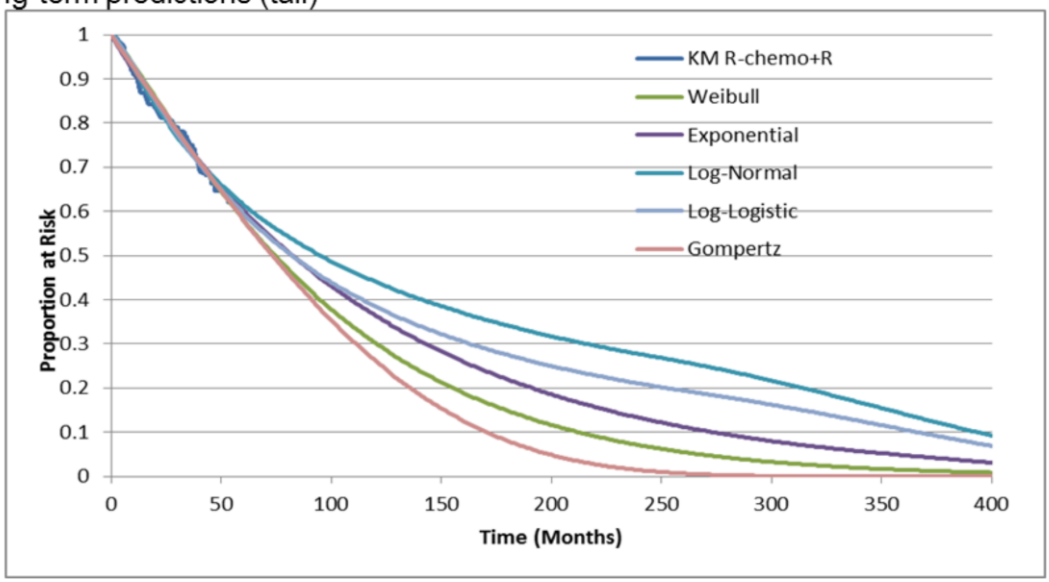
AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion

59

Source: Table 63 in the Company's submission

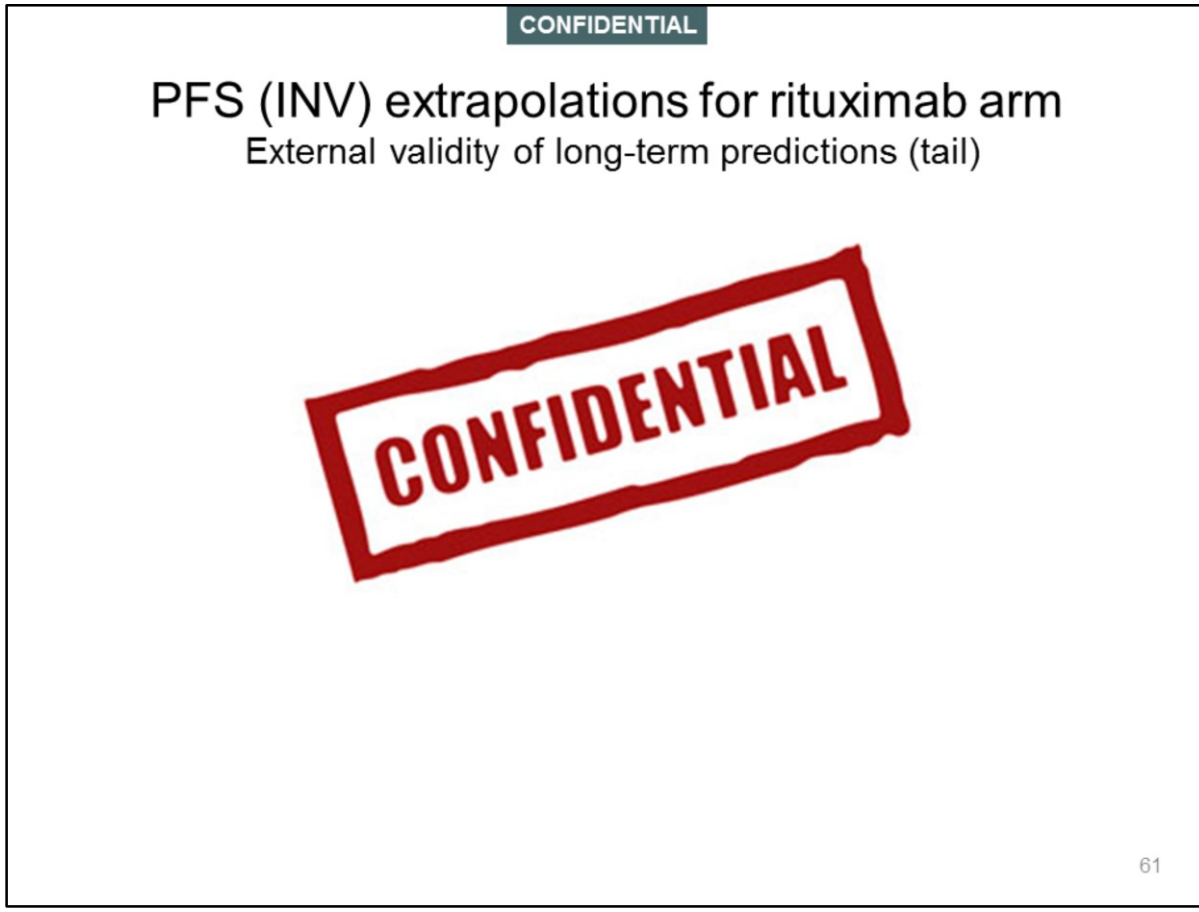
Extrapolating PFS (INV) for rituximab arm

Company chose exponential distribution despite poorer statistical fit to observed data compared with other distributions because it better reflects of long-term predictions (tail)



60

Source: Company's submission Figure 25 (page 143)



Source: Company's submission Figure 26 (page 145)

Predicted progression free survival by parametric curves for rituximab arm (company's model)

	PFS at 6 yrs (%)	PFS at 8 yrs (%)	PFS at 10 yrs (%)	PFS at 15 yrs (%)
Exponential	54.6	44.6	**36.4**	22.0
Weibull	51.3	39.6	30.2	14.9
Log-logistic	54.1	45.2	38.5	27.5
Log-normal	57.1	49.8	44.1	34.2
Generalized Gamma	56.8	49.3	43.5	33.3
Gompertz	50.8	37.4	26.2	8.1

PFS = progression free survival (investigator assessed)

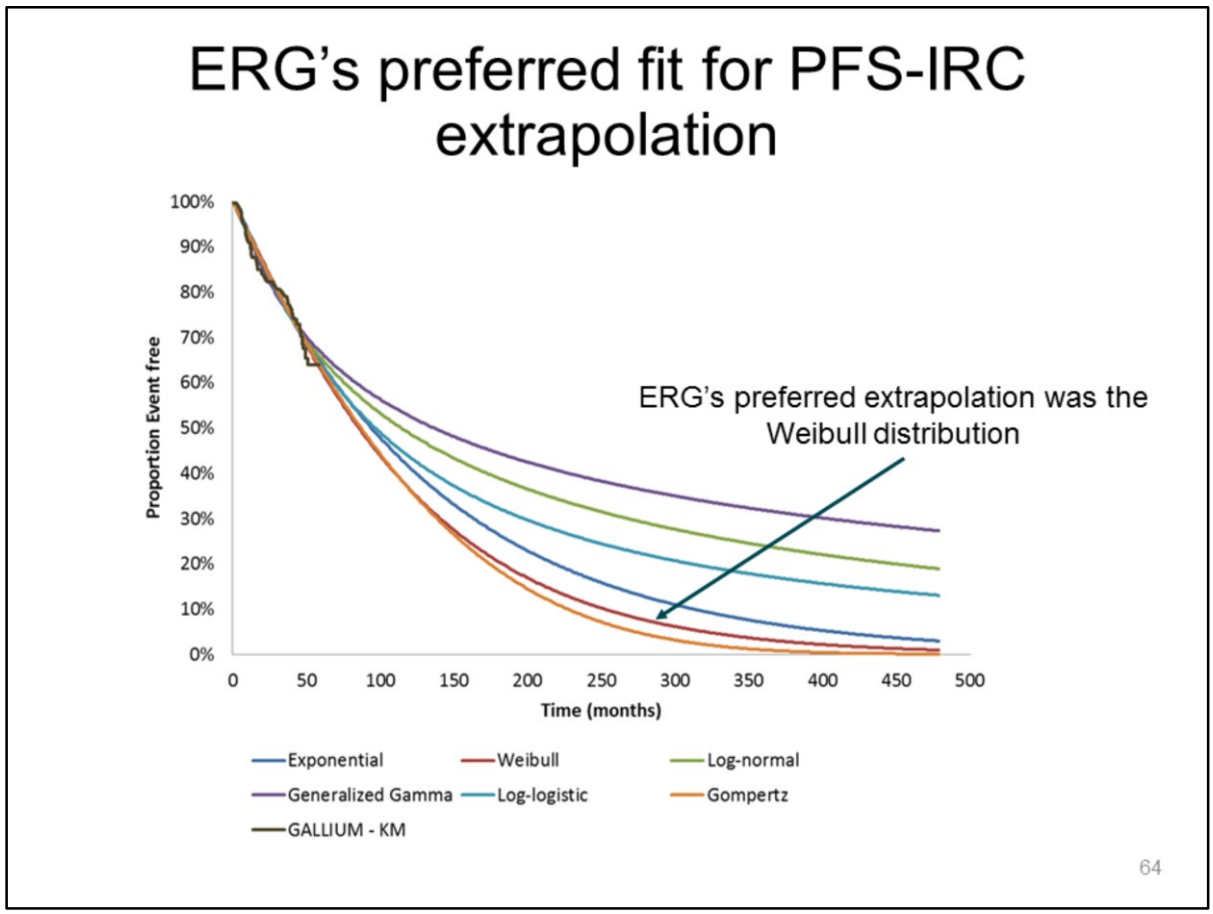
62

Source: Table 64 in the Company's submission

ERG's comments on how company modelled progression-free survival

- Company did "not properly justify" choice of PFS curve
  - difference between exponential and log-logistic unclear
- ERG selected Weibull (predicted PFS at 10 years: 36.4%) for its preferred base case, although Gompertz (predicted PFS at 10 years, 36.3%) also an option
- ERG prefers Weibull because it also fits to the PFS determined by investigator (predicted PFS at 10 years, 30.2%)
- Company stated exponential curve is more conservative and so prefers it over log-logistic
  - But, same reasoning valid for Weibull (preferred by ERG)
- PFS based on independent-review committee less prone to bias (and more conservative) than investigator-assessed PFS; company should use independently reviewed PFS

63



Source Factual accuracy check by the company

Predicted progression free survival by parametric curves for rituximab arm (PFS-IRC)

	PFS at 6 yrs (%)	PFS at 8 yrs (%)	PFS at 10 yrs (%)	PFS at 15 yrs (%)
Exponential	58.8	49.2	41.2	26.5
Weibull	56.4	45.4	**36.4**	20.5
Log-logistic	58.6	50.1	43.5	32.3
Log-normal	61.1	54.2	48.7	38.9
Generalized Gamma	62.9	57.1	52.5	44.4
Gompertz	57.1	45.9	**36.3**	18.6

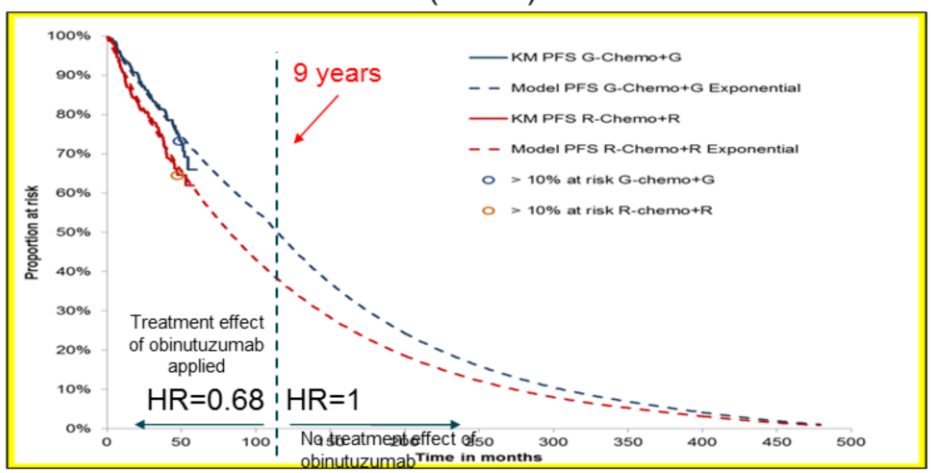
PFS = progression free survival (independent review committee)

65

Source: amended model submitted at clarification stage, Table 5.11 ERG report page79)

Duration of treatment effect on progression-free survival

- **For Extrapolating long term PFS for obinutuzumab arm company**
  - applied HR for PFS to rituximab arm
  - assumed it remained unchanged (did not wane) for 9 years (based on PRIMA)
  - no treatment effect after that (HR=1)



66

Source: Company's submission Figure 27 (page 147)

ERG's comments on duration of effect

- ERG considers a 9 year fixed treatment effect 'speculative'
  - PRIMA compared rituximab maintenance with observation
  - Unclear if one can extrapolate to the duration of treatment effect of obinutuzumab over rituximab
- ERG explored changing the duration of a treatment effect:
  - reasonable to assume a 5-year treatment effect as this is the longest follow up duration from GALLIUM (34 months)

67

Company's approach to modelling death

- Probability of dying before progression: Higher of:
  - Death rate before disease progression from GALLIUM
  - UK age-specific all-cause mortality
- Probability of dying after progression
  - During 'early' progressive state: Higher of:
    - UK age-specific all-cause mortality and
    - Pooled-over-both-arms monthly death rate from GALLIUM
  - During late progressive state: Higher of
    - UK age-specific all-cause mortality
    - Pooled-over-both-arms monthly death rates from people who progressed after 2 years in PRIMA

68

CONFIDENTIAL

Monthly mortality in the model

	N	Events	Monthly (95%CI)
Death rate during progression free-survival state			
Pooled	**1202**	**38**	**0.096% (used in base-case)**
Obin-chemo	601	23	0.113%
R-chemo	601	15	0.078%
Death rates during early progression stage			
Pooled	**155**	**58**	**1.61% (used in base-case)**
Obin-chemo	57	19	1.45%
R-chemo	98	39	1.72%
Death rate during late progressive stage			
PRIMA (late)			[redacted]
PRIMA (early +late)			[redacted]

69

Source: Table 6 and 7 in the clarification response CS.

CONFIDENTIAL

Company's monthly mortality rate

- Post progression survival Kaplan-Meier curves

	GALLIUM	PRIMA
Early progression (<2yrs)	1.61% (base case)	[redacted]
Late progression (>2yrs)	-	[redacted]

70

Source: Appendix 6 of Company's submission, Figure 5 (page 43)

Modelled overall survival

- Calculated as the sum of:
  - Time spent in progression free-survival state
  - Time spent in early progressed disease state
  - Time spent in late progressed disease state
- Modelled overall survival (company's base-case) undiscounted

	O-chemo+O	R-chemo+R	Difference
Mean LY in PFS	11.60	9.68	1.92
Median PFS	9.58	6.83	2.75
Total Mean LY (OS)	19.42	17.97	1.45
Median OS	18.67	16.50	2.17

71

Based on the company's submission table 86, page 189

ERG's comments on company's modelled mortality rates

- ERG questioned pooling of the deaths between 2 arms and preferred separate mortality rate for obinutuzumab and rituximab
- In GALLIUM trial, death rate were higher
  - in obinutuzumab arm during progression free stage
  - In rituximab during early progression stage
- In GALLIUM, no death occurred in patients having late progression

72

Modelling of adverse events

- Company included adverse event costs, but not disutility
- Included only adverse events that affect more than 2% patients in GALLIUM trial
  - But, treated different grades of the same adverse event (for example grade 3, 4 and 5) as separate categories
  - only considered the specific grades of adverse events which affected more than 2% patients

ERG's comments:

- 2% threshold arbitrary
- Considering different grades of the same adverse event separately causes illogical situations (e.g. grade 3 pneumonia was included but grade 4/5 pneumonia were excluded)
- ERG prefer to apply 2% threshold to the pooled grade 3/4/5 adverse events but didn't do it because of 'data and time limitations'
- ERG used same adverse events **list** but counted all of grades 3/4/5 adverse events & also incorporated disutility

73

Company's cost of adverse events in base case

Event (Grade)	Unit Cost	Reference
Anaemia (3)	£2,117	SA03G (NL)
Febrile Neutropenia (3)	£6,226	NICE CG NHL, 2016
Dyspnea (3)	£0	Not costed
Infusion related reaction (3)	£601	SA31E (NS)
Infusion related reaction (4)	£601	SA31E (NS)
Neutropenia (3)	£867	LRiG estimate rev. TA162, TA175
Neutropenia (4)	£867	LRiG estimate rev. TA162, TA175
Pneumonia (3)	£4,155	DZ11P (NL)
Leukopenia (3)	£3,236	SA31E (NL)
Leukopenia (4)	£3,236	SA31E (NL)
Thrombocytopenia (3)	£3,236	SA31E (NL)
Thrombocytopenia (4)	£3,236	SA31E (NL)

Source: Based on Table 82 in the company's submission
***NHS reference costs 2015-16; NL, non-elective long stay; NS, non-elective short stay**

AE disutility used in company's scenario

	Disutility	SE	Source	Duration of adverse event (days)	Source
Neutropenia	-0.09	0.02	Nafees et al., 2008	15.10	NICE TA 306
Thrombocytopenia	-0.11	0.02*	Tolley et al., 2013	23.20	NICE TA 306
Anaemia	-0.12	0.02	Swinburn et al., 2010	16.07	NICE TA 306
Leukopenia	-0.12	0.02	Assumed to be same as Anaemia	16.07	Assumption
Pneumonia	-0.20	0.02	Beusterien et al., 2010	14.00	NICE TA 306

75

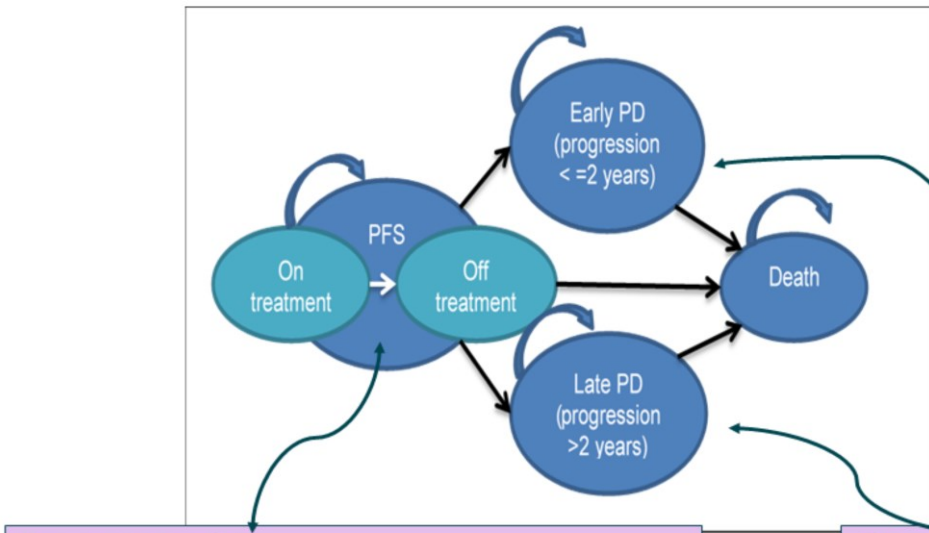
Source: Company's submission table 72

Modelling of health-related quality of life

- EQ-5D data were collected in GALLIUM at baseline, during treatment, after treatment, at the last assessment prior to progression, and the first assessment after progression
- Utilities from GALLIUM used to inform progression-free health state
- Same utility were assumed for both intervention and the comparator arm
- Utility for post-progression state was sourced from literature

76

Company source of utility in model

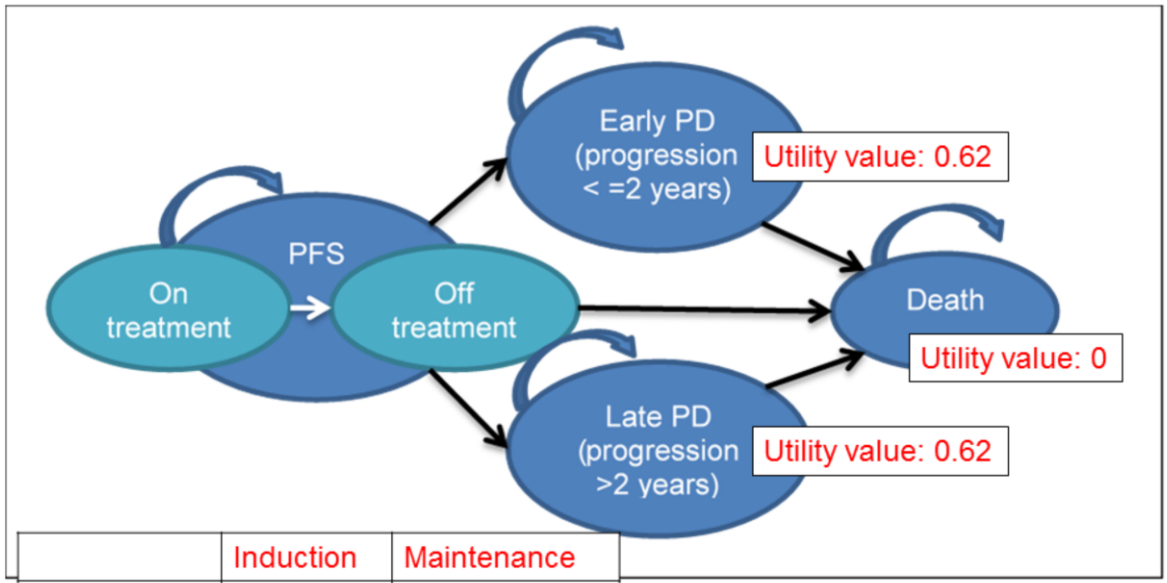


Pre-progression
- EQ-5D from GALLIUM
- Company assumed same utility for both intervention and the comparator arm

Post progression
LITERATURE
Wild et al.

77

Summary: utility values used by company in the model



	Induction	Maintenance
On treatment	0.823	0.823
Off treatment	0.772	0.818

78

Source: Table 73 in the company's submission

ERG's comment on utility values

Unable to verify values

- Company's submission not clear but it seems all available EQ-5D from GALLIUM were used regardless of geographical regions.
- Not mentioned which tariff has been applied to the EQ-5D data in the GALLIUM trial
- Could not find full paper of Wild et al – so could not verify utilities
- Other sources and techniques identified; but utility value for PD health state is "non-transparent and non-replicable"

Health state	Wild et al	Bec et al	GALLIUM	GADOLIN	Mapping FACT-Lym
PFS (on treatment)	0.81	0.71	0.82	0.82	NA
PFS (off treatment)			0.77	0.81	NA
Progressed disease	0.62	0.51	0.78 (early PD) 0.81(late PD)	0.76	0.73

79

For further details see ERG report page 82

ERG's comment on utility values

- For progression-free survival utility values:
  - Unclear whether a UK tariff was applied to GALLIUM estimates
  - Preferred different utility values for treatment arms as per GALLIUM trial, however due to 'time constraints' did not include its base-case
  - Not expected to have a major impact on utilities
- ERG does not agree with company not to adjust utility by decline in age
  - Company: "an age dependent decline is not observed" in trial
  - ERG: after seeing age distribution in GALLIUM, unlikely trial was powered to detect difference in utility between age
- Adverse events were more frequent in the obinutuzumab arm
  - But, estimated utility values higher in obinutuzumab arm vs rituximab "unexpected"
  - This is not reflected in the company's approach were utility values were pooled and should have been incorporated in the base-case
  - ERG considered that disutility should be incorporated

80

CONFIDENTIAL

Resource use and costs
Drug and administration cost

- Drug acquisition and administration costs converted into monthly costs and applied to the **PFS (on-treatment) health state**
- For all drugs, average doses used in GALLIUM were used in base case
- The company has agreed a patient access scheme with DH that makes obinutuzumab available with a confidential discounted price
- Rituximab is also available [redacted]
  - Note: rituximab biosimilars are also available *(see confidential appendix for results that use price of biosimilar)*
- Drug costs for chemotherapy were from the British National Formulary (British National Formulary, 2017) or eMIT (Department of Health, 2016)
- Drug administration costs in the model are based on NHS references costs tariffs (NHS Schedule of Reference Costs 2016) and also included pharmacy cost for IV infusion preparation (£11.50) and patient transport cost (30% patients assumed to need transport costing £39.42)
- For rituximab maintenance, the company assumed that [redacted]% patient would receive it subcutaneously (SC)
  - SC administration cost £227 (IV administration £337)

81

Resource use and costs
Supportive care cost

- Supportive care cost applied to PFS (on and off treatment) early and late progressed disease
- Costs for subsequent treatments incorporated in early and late PD

Supportive care cost

- Haematologist appointment based on ESMO guideline: each visit costs £116 (NHS reference price 2016)
  - monthly during induction (0-6 months),
  - every 3 months 6-30 months in PFS or in early PD
  - every 4 months (30 until progression) or in late PD
  - Diagnostic test/examination: £65.27 with each visit: based on literature
  - CT scan (£132): once during induction, one between 6 to 24 months

82

ESMO: EUROPEAN SOCIETY FOR MEDICAL ONCOLOGY

Resource use and costs
Subsequent treatment cost

- Based on clinical feedback the company assumed that
  - post progression treatment would be the same for both arms
  - costs and outcomes would be similar
  - treatment for early and late progression would not differ
- Data on time to next anti-lymphoma treatment (NALT) from GALLIUM, considered immature and heavily censored → was not used in the base-case
- The company used a total of £13,427 estimated from the cost of subsequent treatment estimated based literature (Papaioannou et al. 2012 for TA110)
- Company considered subsequent costs conservative as they were based on the average costs for all patients and not only those progressing
- In sensitivity analysis the company used costs based on time to next anti-lymphoma treatment data from GALLIUM (average costs of subsequent treatment £5,437.61).

83

ERG's comments on costs

- For bendamustine dose used in the model was 90mg/m^2 which is lower than usual dose of 120mg/m^2
  - Unlikely to have an impact on the results because bendamustine is inexpensive and incorporated in both arms
- ERG critiqued the style of referencing in the company's submission which is not always transparent

84

CONFIDENTIAL

Company's base-case results

Technologies	Total Costs (£)	Total LYG	Total QALYs	Inc Costs (£)	Inc LYG	Inc QALYs	ICER (£/QALY)
Deterministic							
obin-chemo+obin	[redacted]	13.33	10.01	[redacted]	0.84	0.78	[redacted]
R-chemo+R	[redacted]	12.49	9.23	-	-	-	-
Probabilistic (based on 1,000 iteration)							
obin-chemo+obin	[redacted]	13.26	9.98	[redacted]	0.82	0.76	[redacted]
R-chemo+R	[redacted]	12.44	9.21	-	-	-	-

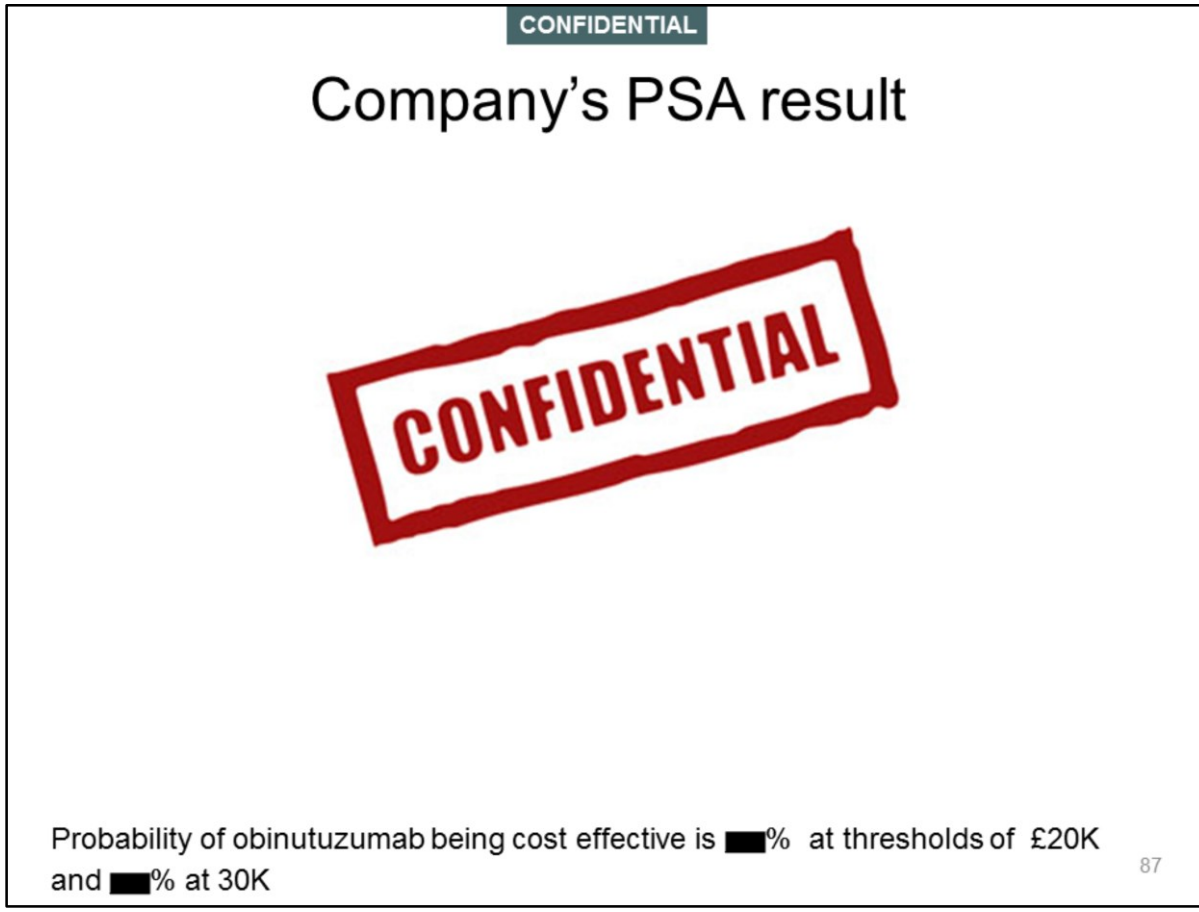
[redacted]

For analyses using acquisition price of biosimilar see Confidential Appendix

85

Source: Based on Table 1 Appendix A – clarification response and electronic model included in the clarification response

Source: Figure 5.12 of ERG report (page, 128)



Source: Figure 5.13 of ERG report (page 128)

For details of model inputs varied during deterministic sensitivity analysis, see ERG report table 5.22
Source: Figure 3 in Appendix A - clarification response

Company's deterministic sensitivity analyses

- Sensitivity analysis showed that the results were most sensitive to the following inputs/assumptions
  1. Duration of treatment effect
  2. Discount rate
  3. PFS extrapolation
  4. Post progression survival
  5. Utility value for progressive disease state

89

Company's scenario analysis

1. Alternative PFS and PPS assumption
   - Log-normal PFS extrapolation with no limit on duration of effect
   - No distinction between EPD and LPD (probability of death equal and based on pooled mortality in PRIMA)
2. Equal QALYs and cost post progression
3. Baseline age based on the median age in Haematological Malignancy Research Network (HMRN) database (63.7 years) instead of GALLIUM (59.0 years)
4. Chemotherapy distribution as per Market Share (base-case as per GALIUM)
5. Separate mortality rate between arms in PFS and Early progressed disease states

90

CONFIDENTIAL

Company's scenario results

	Obinutuzumab		Rituximab				
Scenario	Total cost	Total QALY	Total cost	Total QALY	Δ cost	Δ QALY	ICER
Base-case	[redacted]	12.49	[redacted]	9.23	[redacted]	0.78	[redacted]
1	[redacted]	10.07	[redacted]	9.33	[redacted]	0.74	[redacted]
2	[redacted]	7.20	[redacted]	6.13	[redacted]	1.07	[redacted]
3	[redacted]	9.63	[redacted]	8.92	[redacted]	0.71	[redacted]
4	[redacted]	10.01	[redacted]	9.23	[redacted]	0.78	[redacted]
5	[redacted]	9.98	[redacted]	9.26	[redacted]	0.78	[redacted]

91

ERG's exploratory analyses

- ERG's described its exploratory analysis under 3 heading
  - Fixing errors (correcting the model where the company's electronic model was unequivocally wrong)
  - Fixing violations (correcting the model where the ERG considered that the NICE reference case, scope or best practice has not been adhered to)
  - Matters of judgement (amending the model where the ERG considers that reasonable alternative assumptions are preferred)

Note: [redacted] A biosimilar of rituximab (Truxima, Napp) is available and has a confidential discount. Analyses using the cost of the biosimilar are available in the Confidential Appendix.
All ERG analyses include a minor error relating to administration costs identified during the factual accuracy check by company. These we not corrected because they have a negligible impact on the ICER

92

For details please see the ERG report page 102-104.

CONFIDENTIAL

ERG's exploratory analyses

Scenarios	Obin-chemo+obin		R-chemo+R		Inc Costs (£)	Inc QALYs	ICER (£)
	Total Costs (£)	Total QALYs	Total Costs (£)	Total QALYs			
0. Company's base-case	[redacted]	10.01	[redacted]	9.23	[redacted]	0.78	[redacted]
1. Fixing errors	[redacted]	10.01	[redacted]	9.23	[redacted]	0.78	[redacted]
[redacted]							
2a increasing baseline age	[redacted]	9.63	[redacted]	8.92	[redacted]	0.71	[redacted]
2b chemotherapy distribution	[redacted]	10.01	[redacted]	9.23	[redacted]	0.78	[redacted]
2c female distribution	[redacted]	10.02	[redacted]	9.24	[redacted]	0.79	[redacted]
3 different mortality per arm	[redacted]	9.98	[redacted]	9.26	[redacted]	0.72	[redacted]

93

Source Table 6.1 ERG report (page 116)

CONFIDENTIAL

ERG's Exploratory analyses

Scenarios	Obin-chemo+obin		R-chemo+R		Inc Costs (£)	Inc QALYs	ICER (£)
	Total Costs (£)	Total QALYs	Total Costs (£)	Total QALYs			
0. Company's base-case	[redacted]	10.01	[redacted]	9.23	[redacted]	0.78	[redacted]
Rest of the scenario include scenario 1							
4 age utility decrement	[redacted]	9.48	[redacted]	8.76	[redacted]	0.72	[redacted]
5 PFS-IRC Weibull	[redacted]	10.00	[redacted]	9.34	[redacted]	0.66	[redacted]
6 AE disutility	[redacted]	10.01	[redacted]	9.23	[redacted]	0.78	[redacted]
7 no vial sharing	[redacted]	10.01	[redacted]	9.23	[redacted]	0.78	[redacted]
8 AE grade ≥3 costs & disutilities	[redacted]	10.01	[redacted]	9.23	[redacted]	0.78	[redacted]
(1 to 8 all): ERG preferred base-case	[redacted]	9.12	[redacted]	8.58	[redacted]	0.53	[redacted]

94

Source Table 6.1 ERG report (page 116)

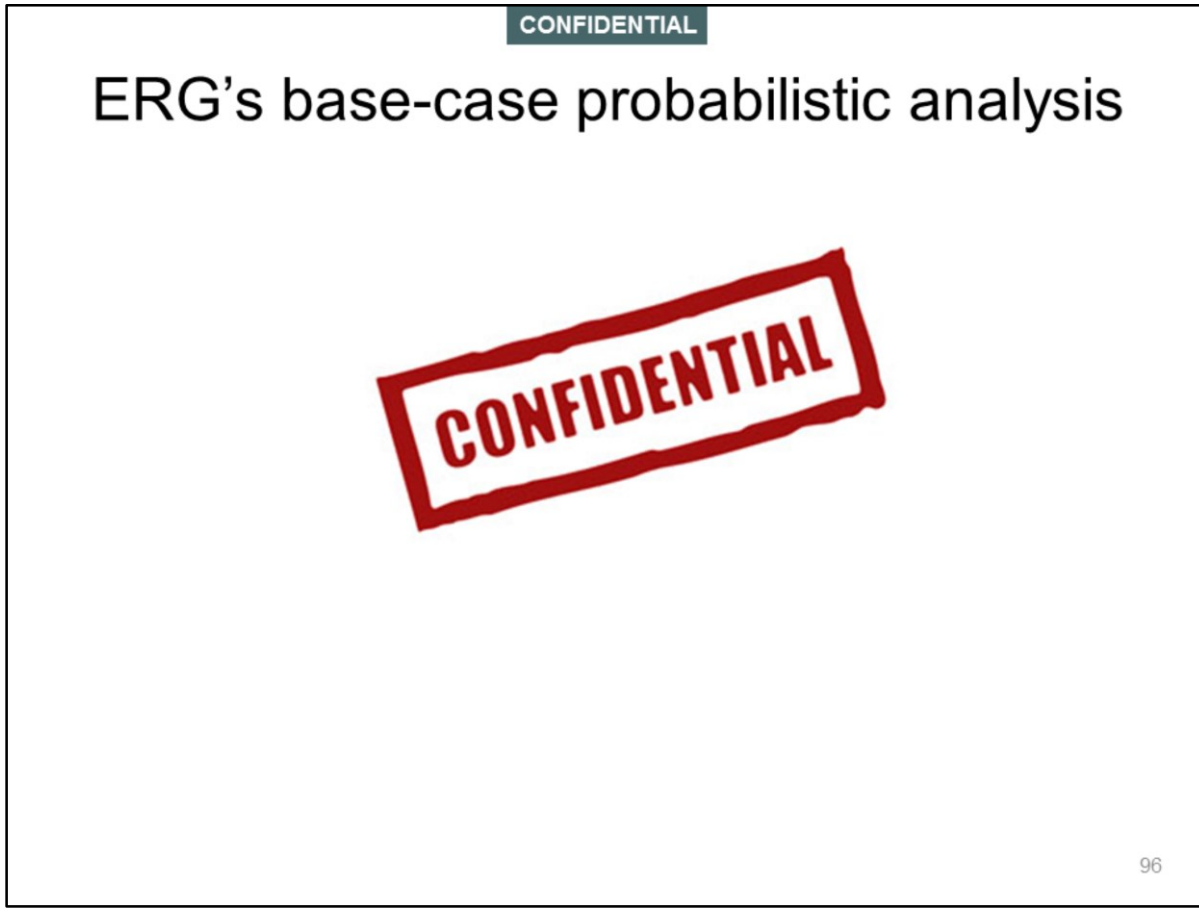
CONFIDENTIAL

ERG's base-case results

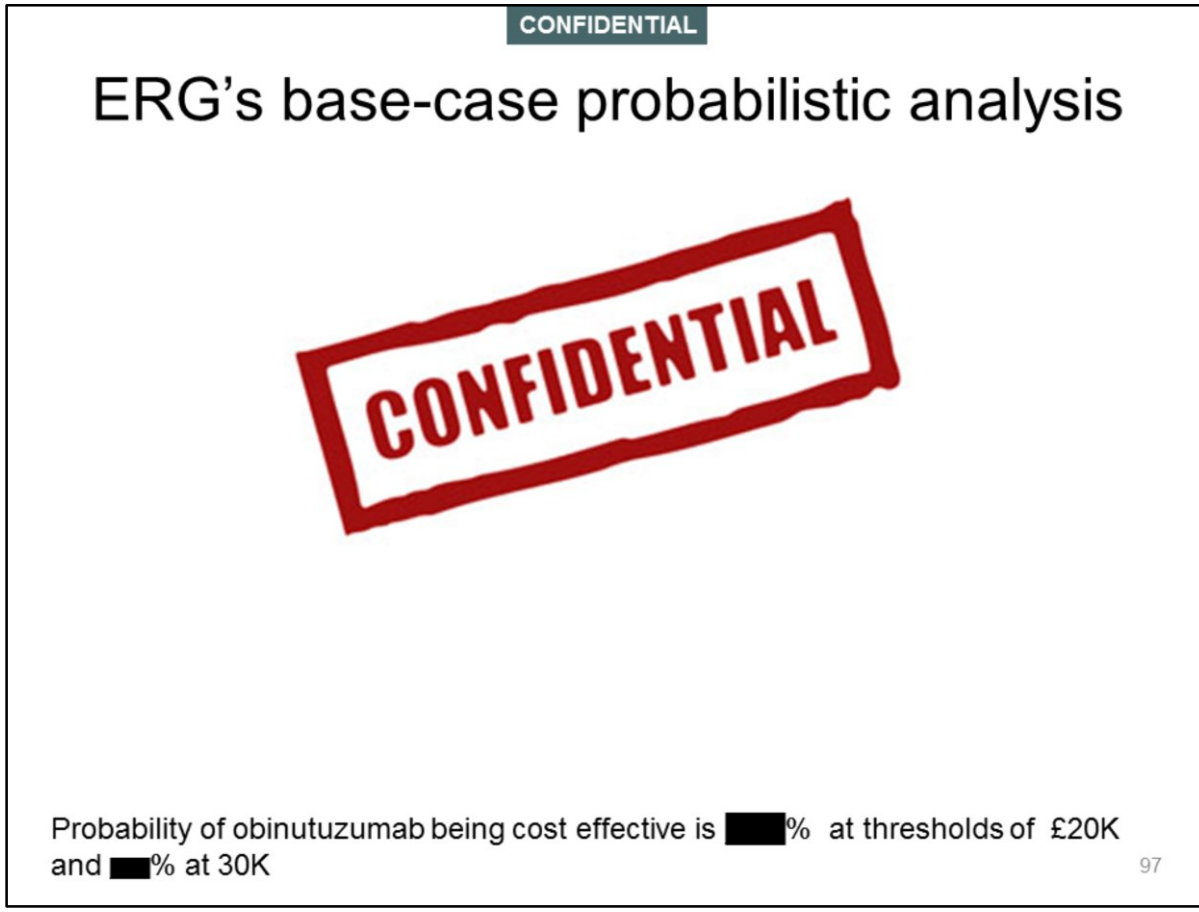
Technologies	Total Costs (£)	Total LYG	Total QALYs	Inc Costs (£)	Inc LYG	Inc QALYs	ICER (£/QALY)
Deterministic							
obin-chemo+obin	[redacted]	12.79	9.12	[redacted]	0.58	0.53	[redacted]
R-chemo+R	[redacted]	12.22	8.58				
Probabilistic (based on 1,000 iteration)							
obin-chemo+obin	[redacted]	12.71	9.07	[redacted]	0.56	0.52	[redacted]
R-chemo+R	[redacted]	12.15	8.55				

95

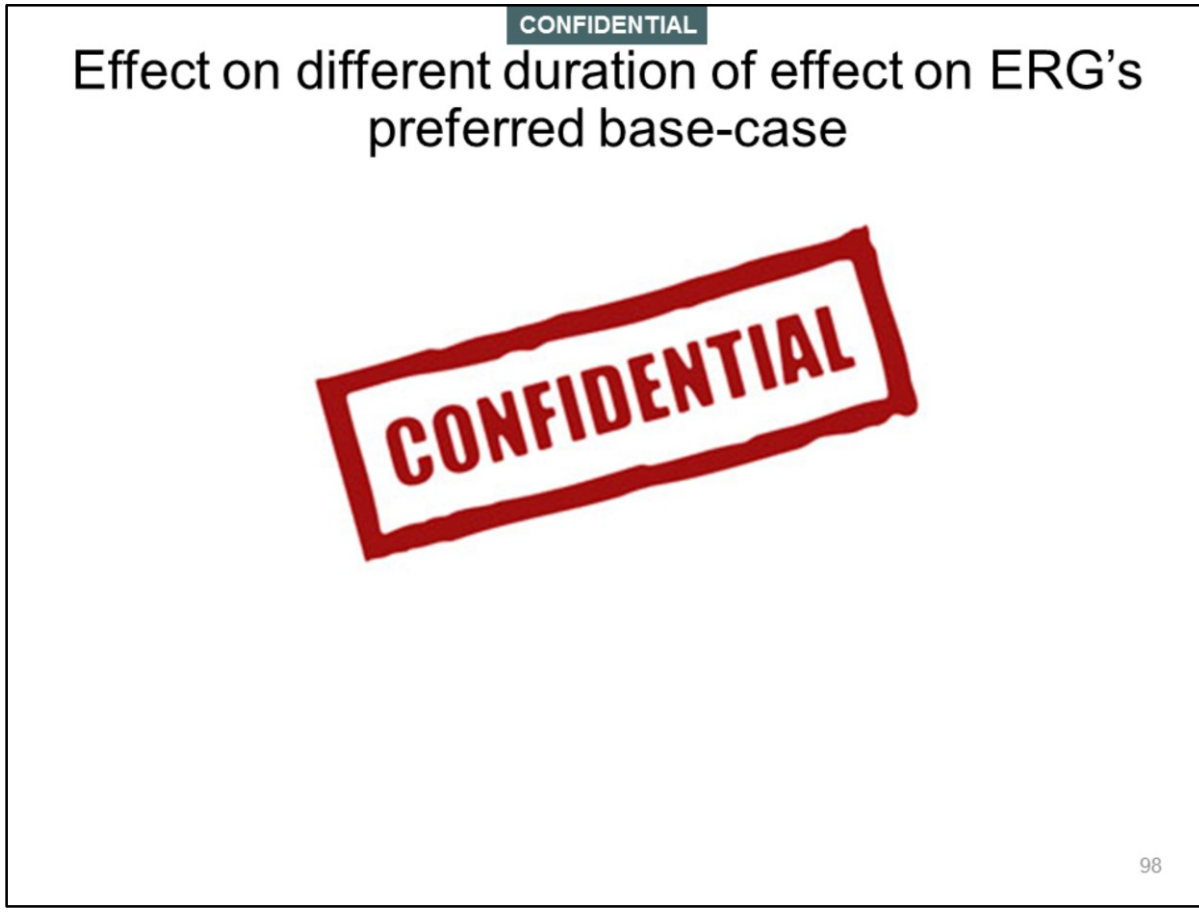
Source: Based on Table 5.30 and 5.33 ERG report



Source: Figure 5.20 ERG report (page 112)



Source: Figure 5.21 ERG report (page 112)



Source: Figure 5.19 ERG report (page 110)

CONFIDENTIAL

ERG's scenario analyses

Scenarios	Inc Costs (£)	Inc QALYs	ICER (£)
Company's base-case	[redacted]	0.78	[redacted]
ERG preferred base-case	[redacted]	0.53	[redacted]
Scenario 1a - (treatment effect duration 5 years)	[redacted]	0.39	[redacted]
Scenario 1b - (PFS Gompertz distribution)	[redacted]	0.52	[redacted]
Scenario 1c - (PFS-INV data)	[redacted]	0.61	[redacted]
Scenario 1d - (pooled mortality)	[redacted]	0.56	[redacted]
Scenario 2a - (no AE disutility)	[redacted]	0.53	[redacted]
Scenario 2b - (no utility decrement with age)	[redacted]	0.58	[redacted]
Scenario 2c - (GALLIUM utilities for PFS and PD health states)	[redacted]	0.44	[redacted]
Scenario 2d - (Wild et al. utilities for PFS and PD health states)	[redacted]	0.52	[redacted]

99

Source: Table 5.34 ERG report (page 109)

CONFIDENTIAL

ERG's scenario analyses

Scenarios	Inc Costs (£)	Inc QALYs	ICER (£)
CS base-case	[redacted]	0.78	[redacted]
ERG preferred base-case	[redacted]	0.53	[redacted]
Scenario 2e - (Bec et al. utilities for PFS and PD health states)	[redacted]	0.47	[redacted]
Scenario 2f - (GADOLIN utilities for PFS and PD health states)	[redacted]	0.45	[redacted]
Scenario 2g - (GALLIUM utilities for PFS and mapping FACT-Lym for PD)	[redacted]	0.47	[redacted]
Scenario 2h - (GALLIUM utilities for PFS and GADOLIN for PD)	[redacted]	0.46	[redacted]
Scenario 2i - (GALLIUM utilities for PFS and Bec et al. for PD)	[redacted]	0.59	[redacted]
Scenario 2j - (Different utilities for early and late PD)	[redacted]	0.53	[redacted]

100

Source: Table 5.34 ERG report (page 109)

CONFIDENTIAL

ERG's scenario analyses

Scenarios	Inc Costs (£)	Inc QALYs	ICER (£)
CS base-case	[redacted]	0.78	[redacted]
ERG preferred base-case	[redacted]	0.53	[redacted]
Scenario 3a - (demographic characteristics in the GALLIUM trial)	[redacted]	0.54	[redacted]
Scenario 4a - (chemotherapy distribution UK market research)	[redacted]	0.53	[redacted]
Scenario 4b - (chemotherapy distribution in the GALLIUM trial – all patients)	[redacted]	0.53	[redacted]
Scenario 4c - (chemotherapy distribution 100% bendamustine)	[redacted]	0.53	[redacted]
Scenario 4d - (chemotherapy distribution 100% CHOP)	[redacted]	0.53	[redacted]
Scenario 4e - (chemotherapy distribution 100% CVP)	[redacted]	0.53	[redacted]
Scenario 4f - (vial sharing)	[redacted]	0.53	[redacted]

101

Source: Table 5.34 ERG report (page 109)

End of life

- Company did not make a case for end-of-life because patients with follicular lymphoma have a life expectancy beyond 24 months

102

Equality

- No equality issues have been identified during scoping or evidence submission

103

Innovation

- The company considered obinutuzumab innovative because
  - a first-in-class Type II glycoengineered anti-CD20 antibody
  - enhanced antibody dependent cellular cytotoxicity, increased direct cell death, and
  - a lower degree of complement dependent cytotoxicity compared with non-glycoengineered, Type I antibodies such as rituximab
  - a meaningful improvement in PFS over rituximab
  - potential to address 'the significant unmet need for this patient population which will provide a significant positive impact on patients' lives'.

104

Authors

- **Anwar Jilani**
  Technical Lead
- **Jasdeep Hayre**
  Technical Adviser
- with input from the Lead Team (Sanjeev Patel, Nigel Westwood, Mark Chapman and Amanda Adler)

105

GALLIUM trial (results)

	Updated analysis (cut-off September 2016)	
Event-Free Survival (event= death, progression, relapse or next anti-lymphoma treatment)		
Patients w/ event, n (%)	130 (21.6%)	179 (29.8%)
HR (stratified), 95% CI	0.66 (0.53 to 0.83)	
Disease-Free Survival (from complete response to progression or death)		
Patients incl. in analysis, n	307	293
Patients w/ event, n (%)	34 (11.1%)	40 (13.7%)

106

GALLIUM trial (results)

	Updated analysis (cut-off September 2016)	
	Obin-chemo n=601	R-Chemo n=601
New Anti-Lymphoma Treatment (non-protocol)		
N (%)	86 (14.3%)	120 (20.0%)
HR (stratified), 95% CI		0.68 (0.52 to 0.90)
Duration of response (from complete/partial response to progression/relapse)		
Patients incl. in analysis, n	569	566
Patients w/ event, n (%)	105 (18.5%)	141 (24.9%)
HR (stratified), 95% CI		0.69 (0.53 to 0.88)

107

CONFIDENTIAL

GALLIUM trial (results)

	Updated analysis (cut-off September 2016)	
	Obin-chemo n=601	R-Chemo n=601
Overall response (CR, PR) at end-of-induction		
Without PET (positron emission tomography) , n (%)	530 (88.2)	519 (86.4)
Δ 95% CI	1.8% (-2.02 to 5.68)	
With PET	N=297	N=298
n (%)	254 (85.5)	242 (81.2)
Δ 95% CI	4.3% (-1.8 to 10.5)	
Complete response at end-of-induction		
Without PET, n (%)	112 (18.6%)	145 (24.1%)
Δ 95% CI	-5.5% (-10.2 to -0.78), p=0.02	
With PET	N=297	N=298
n (%)	184 (62.0%)	169 (56.7%)
Δ 95% CI	5.2% (-2.8 to 13.3), p=0.32	

CONFIDENTIAL

Company's base-case Disaggregated QALY and cost

State	O-chemo-O	R-chemo-R	Difference
Disaggregated QALY			
Progression free survival	7.20	6.13	1.07
Early PD	0.28	0.42	-0.13
Late PD	2.53	2.69	-0.15
Total	10.01	9.23	0.78
Disaggregated cost			
PFS			
Obinutuzumab	[redacted]	0	[redacted]
Rituximab	0	[redacted]	[redacted]
Chemotherapy	371	365	5
Drug Administration	7,751	6,589	1,162
Adverse Events	1,274	1,037	237
Supportive Care	7,759	6,821	938
PFS Total	[redacted]	[redacted]	[redacted]
Progressive disease			
Supportive care and subsequent treatment costs	10,3101	11,956	-1,646
Total PD & PFS	[redacted]	[redacted]	[redacted]

Source: Based on Table 2 Appendix A – clarification response[64]

Values in the table are discounted and half cycle corrected.

Includes a minor error relating to administration costs identified during the factual accuracy check. These we not corrected because they have a negligible impact on the ICER

CONFIDENTIAL

ERG's base-case Disaggregated QALY and cost

State	O-chemo-O	R-chemo-R	Difference
Disaggregated QALY			
Progression free survival	6.71	5.86	0.85
Early PD	0.24	0.33	-0.08
Late PD	2.17	2.40	-0.23
Total	9.12	8.59	0.53
Disaggregated cost			
PFS			
Obinutuzumab	[redacted]	0	[redacted]
Rituximab	0	[redacted]	[redacted]
Chemotherapy	411	406	5
Drug Administration	7,760	6,426	1,334
Adverse Events	737	576	161
Supportive Care	7,595	6,807	788
PFS Total	[redacted]	[redacted]	[redacted]
Progressive disease			
Supportive care and subsequent treatment costs	9,762	11,455	-1,693
Total PD & PFS	[redacted]	[redacted]	[redacted]

Source: Based on Table 5.31 and 5.32 of the ERG report
Values in the table are discounted and half cycle corrected.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal

Obinutuzumab for untreated advanced follicular lymphoma

Final scope

Remit

To appraise the clinical and cost effectiveness of obinutuzumab within its marketing authorisation for untreated advanced follicular lymphoma.

Background

Lymphomas are cancers of the lymphatic system, which is part of the body's immune system, and involve abnormal production of lymphocytes (a type of white blood cell). They are divided into Hodgkin and non-Hodgkin lymphomas. Non-Hodgkin lymphomas are a heterogeneous group of conditions ranging from 'indolent' (low-grade) to 'aggressive' (high-grade) depending on the rate at which the abnormal lymphocytes divide. Indolent lymphomas are slow growing. Follicular lymphoma, which affects B cells, is the most common type of indolent non-Hodgkin lymphoma[1]. People with follicular lymphoma typically present with painless, swollen lymph nodes in the neck, armpit or groin. Lymphomas are commonly staged I (best prognosis) to IV (worse prognosis). The stage of the lymphoma reflects how many groups of lymph nodes are affected, where they are in the body, and whether other organs such as the bone marrow or liver are affected. More people are diagnosed with advanced (stage III or IV) non-Hodgkin lymphoma than early stage disease (stage I and II): 50% are diagnosed with advanced disease, 29% are diagnosed with early stage disease, and in the remainder of cases the stage at diagnosis is not known[2].

In 2013, approximately 11,400 people were diagnosed with non-Hodgkin lymphoma in England, of whom around 20% had follicular lymphoma[2]. The 1-year and 5-year survival rates for people with follicular lymphoma are 96% and 87%, respectively[3].

Advanced-stage follicular lymphoma will initially be treated with chemotherapy, usually in combination with rituximab, and radiotherapy. NICE technology appraisal guidance 243 recommends rituximab in combination with cyclophosphamide, vincristine and prednisolone (CVP), cyclophosphamide, doxorubicin, vincristine and prednisolone (CHOP), mitoxantrone, chlorambucil and prednisolone (MCP), cyclophosphamide, doxorubicin, etoposide, prednisolone and interferon-α (CHVPi) or chlorambucil as an option for untreated symptomatic stage III and IV follicular lymphoma. For people who do not have symptoms, the NICE clinical guideline for non-Hodgkin lymphoma recommends that rituximab is given alone, although at the time of writing this draft scope rituximab monotherapy does not have a marketing authorisation in the UK for untreated non-Hodgkin lymphoma. In addition, bendamustine (which does not currently have a marketing authorisation in the UK for untreated non-Hodgkin lymphoma) has

been available through the Cancer Drugs Fund, alone or in combination with rituximab, as an option for people with untreated indolent non-Hodgkin lymphoma. For people whose follicular non-Hodgkin lymphoma has responded to first-line induction therapy with rituximab in combination with chemotherapy, NICE technology appraisal guidance 226 recommends rituximab maintenance therapy as an option. People whose disease does not respond to treatment, or relapses after treatment is completed, will usually receive a different combination chemotherapy regimen, with or without rituximab. Stem cell transplantation may also be considered.

The technology

Obinutuzumab (Gazyvaro, Roche Products Limited) is a type II monoclonal antibody which binds to the CD20 cell surface antigen on B cells and causes cell death. It is administered intravenously.

Obinutuzumab does not currently have a marketing authorisation in the UK for untreated advanced follicular lymphoma. It has been studied in clinical trials in combination with chemotherapy as an induction treatment, compared with rituximab in combination with chemotherapy, in adults with untreated, advanced, indolent non-Hodgkin lymphoma (including follicular lymphoma). The clinical trials also assessed maintenance treatment with obinutuzumab or rituximab monotherapy, taken until disease progression or for up to 2 years, for people whose disease responded to induction therapy.

Obinutuzumab in combination with bendamustine, followed by obinutuzumab maintenance therapy, has a marketing authorisation in the UK for treating follicular lymphoma in people whose disease did not respond to, or progressed during or up to 6 months after, treatment with rituximab or a rituximab-containing regimen. A NICE technology appraisal of obinutuzumab in this population is ongoing (ID841).

Intervention(s)	Obinutuzumab in combination with chemotherapy, with or without obinutuzumab maintenance therapy
Population(s)	People with untreated advanced follicular lymphoma
Comparators	• Rituximab monotherapy (does not currently have a marketing authorisation in the UK for this indication) • Rituximab-based chemotherapy, with or without rituximab maintenance treatment • Bendamustine monotherapy (does not currently have a marketing authorisation in the UK for this indication; not appraised by NICE but funded via the CDF)

Outcomes	The outcome measures to be considered include: • overall survival • progression-free survival • overall response rate • adverse effects of treatment • health-related quality of life.
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective. The availability of any patient access schemes for the intervention or comparator technologies will be taken into account. The availability and cost of biosimilar products should be taken into account.
Other considerations	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.
Related NICE recommendations and NICE Pathways	**Related Technology Appraisals**: Rituximab for the first-line treatment of stage III-IV follicular lymphoma (2012) NICE Technology Appraisal 243. Review decision August 2014: static guidance list. Rituximab for the first-line maintenance treatment of follicular non-Hodgkin's lymphoma (2011) NICE Technology Appraisal 226. Review decision August 2014: static guidance list. **Appraisals in development (including suspended appraisals)** Obinutuzumab in combination with bendamustine for treating rituximab-refractory follicular lymphoma NICE technology appraisals guidance [ID841]. Publication date to be confirmed.

	Bendamustine in combination with rituximab for the first-line treatment of indolent non-Hodgkin's lymphoma (suspended appraisal) [ID434]. **Related Guidelines** Non-Hodgkin's lymphoma: diagnosis and management (2016). NICE guideline 52. Review date to be confirmed. Haematological cancers: improving outcomes (2016). NICE guideline 47. Review date to be confirmed. **Related NICE Pathways** Non-Hodgkin's lymphoma (2016) NICE pathway
Related National Policy	Department of Health, NHS Outcomes Framework 2016-2017, Dec 2016. Domains 1, 2, 4 and 5. NHS England, National Cancer Drugs Fund List, Sep 2016. NHS England, Manual for prescribed specialised services 2016-2017, May 2016. Chapters 105 and 106 (specialist cancer services, adults and children). Department of Health, Improving Outcomes: A strategy for cancer, fourth annual report, Dec 2014. Department of Health, Cancer commissioning guidance, Dec 2009.

References

1 Cancer Research UK (2014) Different types of non Hodgkin lymphoma. Accessed September 2016

2 Cancer Research UK (2013) Non Hodgkin lymphoma incidence statistics. Accessed September 2016

3 Cancer Research UK (2004–11) Non Hodgkin lymphoma survival statistics. Accessed September 2016

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal

Obinutuzumab for untreated advanced follicular lymphoma [ID1020]

Matrix of consultees and commentators

Consultees	**Commentators (no right to submit or appeal)**
Company • Roche Products (obinutuzumab) Patient/carer groups • None Professional groups • Association of Cancer Physicians • Cancer Research UK • Royal College of Physicians • Royal College of Radiologists • Royal College of Pathologists Others • Department of Health • NHS England • Welsh Government	General • All Wales Therapeutics and Toxicology Centre • British National Formulary • Department of Health, Social Services and Public Safety for Northern Ireland • Healthcare Improvement Scotland • Welsh Health Specialised Services Committee Possible comparator companies • Accord Healthcare (bendamustine, doxorubicin, etoposide, mitoxantrone) **- (Confidentiality agreement not signed, not participating)** • Allergan (bendamustine, etoposide, prednisolone) -**(Confidentiality agreement not signed, not participating)** • Aspen (chlorambucil) • Bausch & Lomb UK (prednisolone) **- (Confidentiality agreement not signed, not participating)** • Baxter Healthcare (cyclophosphamide, mitoxantrone) **- (Confidentiality agreement not signed, not participating)** • Bayer (prednisolone) - **(Confidentiality agreement not signed, not participating)** • Bristol-Myers Squibb (etoposide) **(Confidentiality agreement not signed, not participating)** • Celltrion (rituximab) **(Confidentiality agreement not signed, not participating)** • Concordia International (prednisolone)

Consultees	Commentators (no right to submit or appeal)
	(Confidentiality agreement not signed, not participating) • Dr. Reddy's Laboratories (bendamustine) **(Confidentiality agreement not signed, not participating)** • Hospira UK (mitoxantrone, vincristine) **(Confidentiality agreement not signed, not participating)** • Intrapharm Laboratories (prednisolone) **(Confidentiality agreement not signed, not participating)** • Janssen-Cilag (doxorubicin) • Logixx Pharma Solutions (prednisolone) **(Confidentiality agreement not signed, not participating)** • Medac GmbH (bendamustine, doxorubicin, etoposide) **(Confidentiality agreement not signed, not participating)** • Merck Sharpe & Dohme (interferon alfa) • Mundipharma (rituximab) **(Confidentiality agreement not signed, not participating)** • Napp Pharmaceuticals (bendamustine) • Pfizer (doxorubicin) **(Confidentiality agreement not signed, not participating)** • Roche Products (interferon alfa, rituximab) • Sandoz (cyclophosphamide) **(Confidentiality agreement not signed, not participating)** • Seacross Pharmaceuticals (doxorubicin) **(Confidentiality agreement not signed, not participating)** • Teva UK (doxorubicin) **(Confidentiality agreement not signed, not participating)** • Wockhardt UK (prednisolone) **(Confidentiality agreement not signed, not participating)** • Zentiva (bendamustine, prednisolone)

Consultees	Commentators (no right to submit or appeal)
	(Confidentiality agreement not signed, not participating) Relevant research groups • Institute of Cancer Research Associated Public Health Groups • None

NICE is committed to promoting equality, eliminating unlawful discrimination and fostering good relations between people who share a protected characteristic and those who do not. Please let us know if we have missed any important organisations from the lists in the matrix, and which organisations we should include that have a particular focus on relevant equality issues.

PTO FOR DEFINITIONS OF CONSULTEES AND COMMENTATORS

Definitions:

Consultees

Organisations that accept an invitation to participate in the appraisal; the company that markets the technology; national professional organisations; national patient organisations; the Department of Health and the Welsh Government and relevant NHS organisations in England.

The company that markets the technology is invited to make an evidence submission, respond to consultations, nominate clinical specialists and has the right to appeal against the Final Appraisal Determination (FAD).

All non-company consultees are invited to submit a statement[1], respond to consultations, nominate clinical specialists or patient experts and have the right to appeal against the Final Appraisal Determination (FAD).

Commentators

Organisations that engage in the appraisal process but that are not asked to prepare an evidence submission or statement, are able to respond to consultations and they receive the FAD for information only, without right of appeal. These organisations are: companies that market comparator technologies; Healthcare Improvement Scotland; other related research groups where appropriate (for example, the Medical Research Council [MRC], National Cancer Research Institute); other groups (for example, the NHS Confederation, NHS Alliance and NHS Commercial Medicines Unit, and the British National Formulary.

All non-company commentators are invited to nominate clinical specialists or patient experts.

[1] Non-company consultees are invited to submit statements relevant to the group they are representing.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Gazyvaro▼ (obinutuzumab) in combination with chemotherapy for the first-line treatment of patients with advanced follicular lymphoma

Company evidence submission

Roche Products Limited

May 2017

File name	Version	Contains confidential information	Date
ID1020 Gazyvaro for first-line FL [ACIC]	**V1**	**Yes**	**10 May 2017**

Contents

Tables

Figures

Abbreviations

ADCC	Antibody-dependent cellular cytotoxicity
ADCP	Antibody-dependent cellular phagocytosis
AE	Adverse event
AESI	Adverse events of special interest
AIC	Akaike Information Criterion
ALT	Alanine transaminase
ANC	Absolute neutrophil count
ASCO	American Society of Clinical Oncology
ASCT	Autologous stem cell transplantation
ASH	American Society of Hematology
AST	Aspartate transaminase
BCSH	British Committee for Standards in Haematology
BIC	Bayesian Information Criterion
BM	Bone marrow
BMI	Body mass index
BNF	British National Formulary
BOR	Best overall response
BSA	Body surface area
CADTH	Canadian Agency for Drugs and Technologies in Health
CDC	Complement dependent cytotoxicity
CDF	Cancer Drugs Fund
CHMP	Committee for Medicinal Products for Human Use
CHOP	Cyclophosphamide, doxorubicin, vincristine and prednisone
CHVP	Cyclophosphamide, etoposide, doxorubicin and prednisolone
CI	Confidence interval
CLL	Chronic lymphocytic leukaemia
CMH	Cochran–Mantel–Haenszel test
COMP	Committee for Orphan Medicinal Products
CR	Complete response
CRD	Centre for Reviews and Dissemination
CSR	Clinical Study Report
CT	Computed tomography
CVP	Cyclophosphamide, vincristine, and prednisone
DFS	Disease-free survival
DLBCL	Diffuse large B cell lymphoma
DSU	Decision Support Unit
ECOG	Eastern Cooperative Oncology Group
EFS	Event-free survival
EHA	European Haematology Association
EMA	European Medicines Association
EOI	End of induction
EOMR	End of maintenance response
EQ-5D	EuroQoL-5 dimension
ESMO	European Society for Medical Oncology
FACT-G	Functional Assessment of Cancer Therapy - General

FACT-Lym	Functional Assessment of Cancer Therapy for Patients with Lymphoma
FACT-TOI	Functional Assessment of Cancer Therapy - Trial Outcome Index
FC	Fludarabine and cyclophosphamide
FDA	Food and Drug Administration
FDG-PET	Fludeoxyglucose positron-emission tomography
FL	Follicular lymphoma
FLIPI	Follicular Lymphoma International Predictive Index
FU	Follow up
G	Gazyvaro (maintenance monotherapy)
G-chemo	Gazyvaro in combination with chemotherapy as induction therapy
G-CSF	Granulocyte colony-stimulating factor
GI	Gastrointestinal
HDT	High-dose therapy
HMRN	Haematological Malignancy Research Network
HR	Hazard ratio
HRQoL	Health-related quality of life
HSUV	Health state utility
ICER	Incremental cost effectiveness ratio
ICML	International Conference on Malignant Lymphoma
ICTRP	International Clinical Trials Registry Platform
IDMC	Independent Data Monitoring Committee
INAHTA	International Network of Agencies for Health Technology Assessment
iNHL	Indolent non-Hodgkin lymphoma
INR	International normalised ratio
IOG	Improving Outcomes Guidance
IRC	Independent review committee
IRR	Infusion related reactions
ISPOR	International Society For Pharmacoeconomics and Outcomes Research
ITT	Intent-to-treat
IVRS	Interactive Voice Response System
KM	Kaplan Meier
LAA	Last antibody administration
LDH	Lactate dehydrogenase
LVEF	Left ventricular ejection fraction
LYG	Life years gained
MALT	Mucosa associated lymphoma tissue
MCP	Mitoxantrone, chlorambucil and prednisolone
MRD	Minimal residual disease
MRI	Magnetic Resonance Imaging
MTX	Methotrexate
MUGA	Multigated radionuclide angiography
MZL	Marginal zone lymphoma
NALT	New anti-lymphoma treatment
NCCN	National Comprehensive Cancer Network
NE	Not estimated
NIH	US National Institute of Health
ORR	Overall response rate

OS	Overall survival
PAS	Patient access scheme
PB	Peripheral Blood
PCR	Polymerase chain reaction
PD	Progressed disease
PET	Positron-emission tomography
PFS	Progression-free survival
PML	Progressive multifocal leukoencephalopathy
PPS	Post-progression survival
PR	Partial response
PRO	Patient reported outcomes
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
PTT	Partial thromboplastin time
QALY	Quality adjusted life year
R	MabThera (Rituximab) (as maintenance monotherapy in GALLIUM)
R-chemo	MabThera in combination with chemotherapy as induction therapy
SACT	Systemic Anti-Cancer Therapy (chemotherapy dataset)
SAE	Serious adverse event
SD	Stable disease
SE	Standard error
SLL	Small lymphocytic lymphoma
SOC	Standard of care
TLS	Tumour lysis syndrome
TTTD	Time-to-treatment-discontinuation
ULN	Upper limit of normal
WHO	World Health Organisation
WM	Waldenstrom macroglobulinaemia

1. Executive summary

1.1 Statement of decision problem

The decision problem (summarised in Table 1 below) in comparison to the final scope issued by NICE, adressess the comparison of Gazyvaro (obinutuzumab) in combination with chemotherapy, followed by Gazyvaro maintenance therapy in patients achieving a response (G-chemo+G) to MabThera (rituximab) in combination with chemotherapy, followed by MabThera maintenance therapy in patients achieving a response (R-chemo+R) in patients with previously untreated advanced follicular lymphoma (FL).

Table 1: The decision problem

	Final scope issued by NICE		**Rationale if different from the final NICE scope**
Population	People with untreated advanced follicular lymphoma	People with untreated advanced follicular lymphoma	No difference
Intervention	Obinutuzumab in combination with chemotherapy, with or without obinutuzumab maintenance therapy	Gazyvaro (obinutuzumab) in combination with chemotherapy (CVP, CHOP or bendamustine), followed by Gazyvaro maintenance therapy in patients achieving a response	Align with wording of expected Marketing Authorisation
Comparator (s)	• Rituximab monotherapy (does not currently have a Marketing Authorisation in the UK for this indication) • Rituximab-based chemotherapy, with or without rituximab maintenance treatment • Bendamustine monotherapy (does not currently have a Marketing Authorisation in the UK for this	• MabThera (rituximab) in combination with chemotherapy, followed by MabThera maintenance therapy in patients achieving a response	• Induction with MabThera monotherapy is not an appropriate comparator for patients with advanced, symptomatic FL for which the standard of care is MabThera in combination with chemotherapy. NICE guidelines recommend the use of MabThera monotherapy induction in advanced asymptomatic patients only who would not be treated with

	indication; not appraised by NICE but funded via the CDF)		chemotherapy but may be managed by observation ('watch and wait'). • Wording on MabThera use aligned with use in current clinical practice • SACT and market research data indicates little use of bendamustine as monotherapy. Bendamustine is considered only in combination with MabThera in the first-line FL induction setting
Outcomes	The outcome measures to be considered include: • overall survival • progression-free survival • overall response rate • adverse effects of treatment • health-related quality of life	The outcome measures to be considered include: • overall survival • progression-free survival • overall response rate • adverse effects of treatment • health-related quality of life	No difference
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective. The availability of any patient access schemes for the intervention or comparator technologies will be taken into account.	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective. The availability of any patient access schemes for the intervention or comparator technologies will be taken into account.	No difference

	The availability and cost of biosimilar products should be taken into account.	The availability and cost of biosimilar products should be taken into account.	
Subgroups to be considered	None	None	No difference
Special considerations including issues related to equity or equality	None identified	None identified	No difference

1.2 Description of the technology being appraised

A summary of the Marketing Authorisation, indications, restrictions and methods of administration for Gazyvaro is presented below.

Table 2: Technology being appraised

UK approved name and brand name	Obinutuzumab (Gazyvaro)
Marketing Authorisation/CE mark status	An application for UK Marketing Authorisation was made for Gazyvaro in combination with chemotherapy, followed by Gazyvaro monotherapy as maintenance on 16th March 2017. Committee for Medicinal Products for Human Use (CHMP) opinion is anticipated in [redacted], with regulatory approval expected in [redacted].
Indications and any restriction(s) as described in the summary of product characteristics	Gazyvaro currently has Marketing Authorisation for the following indications: • In combination with chlorambucil for the treatment of adult patients with previously untreated chronic lymphocytic leukaemia (CLL) and with comorbidities making them unsuitable for full-dose fludarabine based therapy. • Gazyvaro in combination with bendamustine followed by Gazyvaro maintenance for the treatment of patients with follicular lymphoma who did not respond or who progressed during or up to 6 months after treatment with rituximab or a rituximab-containing regimen." As noted in the draft summary of product characteristics (SmPC), Gazyvaro will only be contraindicated to people who demonstrate hypersensitivity to the medicinal product or any of its excipients.
Method of administration and dosage	***Gazyvaro Induction in combination with chemotherapy (G-chemo):*** • With CHOP: 1000 mg fixed dose Gazyvaro on days 1, 8, and 15 of Cycle 1 and on Day 1 of Cycles 2–8 (21-day cycles) • With CVP: 1000 mg fixed dose Gazyvaro on days 1, 8, and 15 of Cycle 1 and on Day 1 of Cycles 2-8 (21-day cycles) • With bendamustine: 1000 mg fixed dose Gazyvaro on days 1, 8, and 15 of Cycle 1 and on Day 1 of Cycles 2–6 (28-day cycles) ***Gazyvaro maintenance:*** 1000 mg fixed dose Gazyvaro once every 2 months

	for up to two years or until progression, whichever occurs first

1.3 Summary of the disease and current clinical practice

Non-Hodgkin lymphoma (NHL) is a heterogeneous group of lymphoproliferative malignancies, of which 80–95% of cases arise from B-cells and the remaining from T-cells (Rancea et al., 2014). Follicular lymphoma (FL) is the second most common NHL diagnosed in Western Europe and the United States, comprising over 35% of all NHLs and 70% of indolent lymphomas (Freedman, 2015). The median age at diagnosis of FL in the UK is approximately 65 years old (Haematological Malignancy Research Network, 2017a).

FL tends to be insidious in nature. Typically, patients present with multiple sites of lymphadenopathy and/or bone marrow disease [advanced-stage disease (III/IV)]. This may manifest itself with disease-related symptoms such as fatigue, weight loss, fever and night sweats (Pettengell et al., 2008).

Patients with advanced stage FL are usually considered incurable with standard therapeutic approaches therefore treatment generally attempts to control the disease. FL is typified by a chronic course comprising of repeated relapses, treatment and progression. Generally, median life expectancy ranges have been reported from 8–12 years after diagnosis, although this has extended to around 15 years in the post-rituximab era (Tan et al., 2013). Strategies to predict survival have been implemented including the Follicular Lymphoma International Predictive Index (FLIPI) (Solal-Celigny et al., 2004). A revised FLIPI, known as the FLIPI2, was developed to separate patients with significantly different hazard ratios for progression/relapse in the era of anti-CD20 monoclonal antibody treatments (Federico et al., 2009).

Despite the indolent nature of FL, approximately 90% of newly diagnosed stage II–IV FL patients eventually require systemic treatment (Nastoupil et al., 2015). Rituximab-containing regimens are standard of care for first-line treatment of patients with advanced, symptomatic FL and maintenance therapy for two years has become accepted practice in the UK. However, approximately 20% of FL patients who receive immunochemotherapy still suffer disease progression within two years from diagnosis (Casulo et al., 2015b) and those patients with early progression have poorer outcomes (Maurer et al., 2016). Patients who experience early progression have a significantly increased risk of death than patients who do not relapse within the first 24 months (Casulo et al., 2015b). Furthermore, studies have shown that some patients with FL who progressed or relapsed after treatment with rituximab

or a rituximab-containing regimen have a very short progression-free survival (PFS) of less than one year (range 5.8–10.4 months) (Czuczman et al., 2012, Friedberg et al., 2008, Horning et al., 2005, Kahl et al., 2010).

1.4 Summary of the clinical effectiveness analysis

Nomenclature used for the GALLIUM study in the clinical effectiveness section:

- Gazyvaro in combination with chemotherapy as induction therapy is abbreviated as **G-chemo**
- Gazyvaro monotherapy as maintenance is abbreviated as **G**
- The complete treatment regimen, i.e. induction plus maintenance therapy is therefore abbreviated as **G-chemo+G. This represents the regimen as per the anticipated Marketing Authorisation**
- MabThera in combination with chemotherapy as induction therapy is abbreviated as **R-chemo**
- MabThera monotherapy as maintenance is abbreviated as **R**
- The complete treatment regimen, i.e. induction plus maintenance therapy is therefore abbreviated as **R-chemo+R.**

Evidence for the clinical effectiveness, safety and tolerability of Gazyvaro in combination with chemotherapy as induction therapy, followed by Gazyvaro monotherapy maintenance has been demonstrated in one Phase III, open-label randomised controlled trial, GALLIUM (BO21223, NCT01332968) (Clinical Trials.Gov).

This study compared Gazyvaro combined with chemotherapy (**G-chemo**) as induction, followed by Gazyvaro monotherapy as maintenance (**G**), with MabThera combined with chemotherapy (**R-chemo**) as induction followed by MabThera monotherapy as maintenance (**R**) in previously untreated patients with advanced iNHL requiring treatment. Of the 1,401 patients randomised in the study, 1,202 were diagnosed with FL; the primary endpoint of GALLIUM was PFS as assessed by the investigator in patients with FL.

The median age of patients was 59.0 years (range: 23–88 years). The overall median time from first diagnosis to randomisation was 1.5 months (range: 0.0–168.1 months). The greatest proportion of patients comprised intermediate and high-risk FLIPI (37.2% and 41.8% respectively) and FLIPI-2 groups (50.3% and 40.6%, respectively), and Ann Arbor stage III—IV (>91%)

This submission presents the results of the primary data analyses (clinical cut-off 31st January 2016), including all 1,401 patients randomised to the two study arms. Data from an updated analysis (clinical cut-off 16th September 2016) is presented where available.

Efficacy in GALLIUM

The primary endpoint was met in GALLIUM as G-chemo+G treatment resulted in a statistically significant increase in investigator-assessed PFS compared with R-chemo+R. G-chemo+G therapy significantly reduced the risk of experiencing a PFS event by 34% compared with R-chemo+R treatment (stratified HR 0.66, 95% CI: 0.51, 0.85; p=0.0012) (F. Hoffmann-La Roche Ltd, 2016).

On the basis of Kaplan-Meier (KM) estimates, 80.9% (95% CI, 77.4%, 84.0%) and 73.3% (95% CI: 68.8, 77.2) of patients in the R-chemo+R arm were progression-free at two and three years, respectively, compared with 87.7% (95% CI, 84.6%, 90.1%) and 80.0% (95% CI: 75.9, 83.6) of patients in the G-chemo+G arm. The results of the independent-review committee (IRC) assessment of PFS were consistent with the investigator-assessed PFS results (stratified hazard ratio [HR] 0.71 [95% CI 0.54; 0.93]; p=0.0138), while other secondary time-to-event endpoints (overall survival, event-free survival, disease-free survival, duration of response, and time to next anti-lymphoma treatment) were supportive of the PFS outcomes.

Reflecting the indolent nature of FL disease, and after a median follow-up of approximately 34.5 months, median PFS was not expected to be reached at interim analysis. Based on the PRIMA study[1], where 59.2% of previously untreated FL patients on R maintenance were progression-free at 6 years after 73 months' median follow-up (Seymour JF et al., 2013), and assuming a conservative median PFS of six years for R-chemo+R, the observed HR of 0.66 in GALLIUM would translate to a 1.5x longer median PFS for G-chemo+G than R-chemo+R, and to an estimated three year improvement in the G-chemo+G arm. Longer follow-up data will confirm if these benefits are achieved.

Several pre-specified subgroup analyses showed that the investigator-assessed PFS benefit with G-chemo+G was consistent across all patient subgroups with the exception of low risk FLIPI (HR 1.17 [95% CI: 0.63, 2.19]; based on 253 patients). Excluding low risk FLIPI, the observed hazard ratios were below 1.00 and ranged from 0.40–0.86 for subgroups including at least 10% of patients. GALLIUM was not designed to compare the three different

[1]The PRIMA Phase III study was designed to investigate the potential benefit of 2-years of R maintenance in patients with FL responding to one of three non-randomised first line immunochemotherapy treatments

chemotherapy regimens used in the study (CHOP, CVP or bendamustine). As the allocation of chemotherapy was not randomised at the patient level, there may be confounding differences in baseline patient characteristics between the chemotherapy subgroups. Pre-planned subgroup analyses of investigator-assessed PFS HRs showed that all G-containing chemotherapy regimens had a consistent benefit over R-chemo regimens in FL patients (CHOP, 0.77 [95% CI, 0.50–1.20]; CVP, 0.63 [95% CI, 0.32–1.21]; bendamustine, 0.61 [95% CI, 0.43–0.86]).

GALLIUM was not designed to evaluate treatment benefits separately for the induction and maintenance phases; however, it is very unlikely that Gazyvaro would provide the PFS benefit observed in GALLIUM if it was used as maintenance only. Minimal residual disease (MRD)-negativity and complete response (CR) rates with positron emission tomography (PET) at the end-of-induction (EOI) were significantly higher in the G-chemo arm, which suggests that Gazyvaro may induce deeper responses than rituximab during induction. Furthermore, a significantly greater proportion of patients in the G-chemo arm achieved MRD-negative status in peripheral blood at mid-induction (94.3% vs. 88.9%; p=0.0132) and in PB and/or bone marrow at the EOI (92.0% vs. 84.9%; p=0.0041) compared with patients in the R-chemo arm. These findings suggest that G-chemo based induction may induce more rapid and more effective tumour-cell clearance than R-chemo based treatment.

Safety in GALLIUM

The toxicity of G-chemo+G was clinically manageable in GALLIUM, as indicated by the high completion rate of dosing and the limited number of dose delays and withdrawals due to adverse events (AEs), which is supported by the similar impact on quality of life between the two treatment arms.

The nature of AEs observed were consistent with the known profiles of the study treatments, with a similar incidence of all grade AEs in the two arms; 98.0% of patients in the R-chemo+R arm vs. 99.5% of patients in the G-chemo+G arm. While patients in the G-chemo+G arm had a numerically higher frequency of grade 3 to 5 AEs and serious AEs than patients in the R-chemo+R arm, the rate of fatal (grade 5) AEs was comparable between the treatment arms. Overall, although the frequency of some AEs was higher in the G-chemo+G arm, no new or unexpected safety signals were detected.

Bendamustine was associated with higher rates of severe infections than CHOP or CVP during maintenance and follow up in both treatment arms. Non-relapse fatal AEs were also more common in bendamustine-treated patients during all study phases, although absolute

numbers were small. The MRD data from GALLIUM provide evidence that less intensive chemotherapy regimens combined with Gazyvaro still demonstrate greater efficacy than when given with rituximab and maintain the overall beneficial effect of Gazyvaro (Pott C et al., 2016).

Relevance to UK clinical practice

The study population in GALLIUM is largely reflective of the advanced FL population in the UK. More patients were recruited from the UK than any other country (293 from 29 centres), indicating that the results of GALLIUM will reflect UK practice. Furthermore, feedback from clinical experts[2] confirms that the baseline characteristics of FL patients enrolled into GALLIUM are reflective of the population seen in UK clinical practice.

Gazyvaro is compared against a relevant active comparator in GALLIUM as R-chemo followed by rituximab maintenance therapy is regarded as the standard of care for the first-line treatment of patients with advanced symptomatic FL. Furthermore, GALLIUM was designed to capture endpoints which are relevant to UK clinical practice and that address the unmet medical need for this patient population.

1.5 Summary of the cost-effectiveness analysis

Nomenclature used for GALLIUM in the cost-effectiveness section:

● Gazyvaro (obinutuzumab; G) or MabThera (rituximab; R) in combination with chemotherapy as induction therapy, followed by G or R monotherapy as maintenance is abbreviated as G-chemo+G and R-chemo+R, respectively. G-chemo+G represents the regimen as per the anticipated Marketing Authorisation.

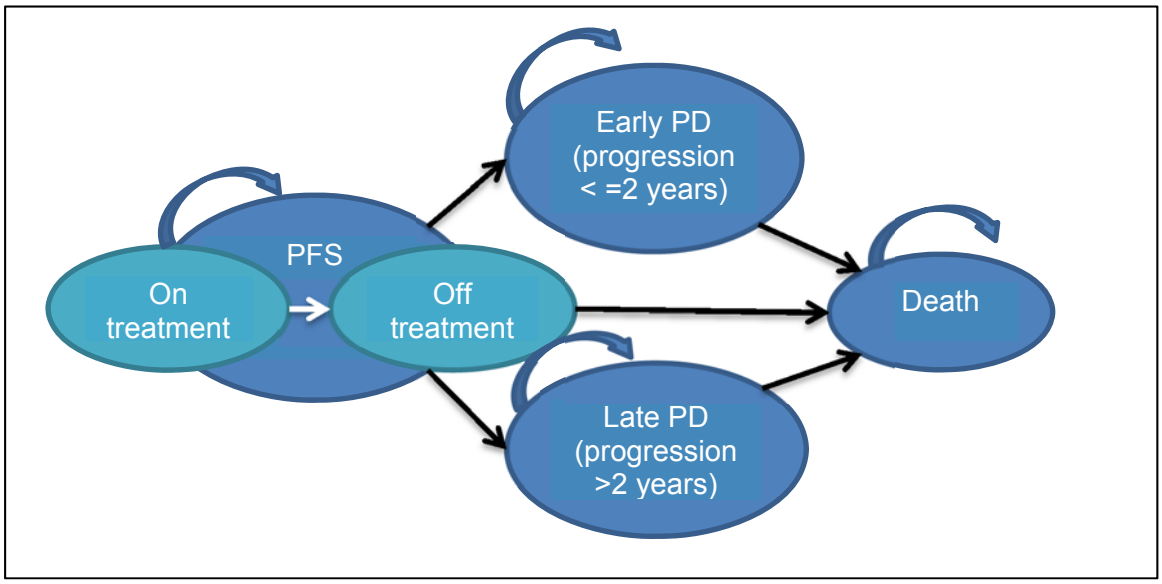
A de novo cost utility analysis was conducted in order to evaluate the cost-effectiveness of Gazyvaro in combination with chemotherapy followed by Gazyvaro monotherapy as maintenance for patients responding to induction (**G-chemo+G**) compared with MabThera in combination with chemotherapy followed by MabThera monotherapy as maintenance for patients responding to induction (**R-chemo+R**) from the perspective of the NHS and personal social care services.

A four state transition Markov model with a one month cycle duration was constructed in Microsoft Excel® to explore the health outcomes and costs associated with G-chemo+G

[2]An expert advisory board was consulted at a one-day meeting in April 2017. The panel consisted of consultant haematologists specialising in the management of patients with FL, many of whom have experience of Gazyvaro from clinical trials. The panel was selected based on their significant clinical and research experience.

compared to R-chemo+R in patients with FL who were previously untreated and required treatment. The model structure is shown in Figure 1 with health states PFS-on or PFS-off treatment, early progressed disease (early PD), late PD and death. Introduction of two PD states allowed modelling of different outcomes and costs for patients progressing within two years of starting first treatment (early PD) versus patients progressing after two years (late PD).

Figure 1: De novo model structure



PD, progressed disease; PFS, progression-free survival

A time horizon of 40 years was used to capture all costs and benefits associated with the treatment with G-chemo+G over a patient's lifetime. Costs and quality adjusted life years (QALYs) were discounted at 3.5%.

Clinical data sources

Patient level data from the GALLIUM trial provided the clinical data inputs for the model to estimate time in PFS and early PD. Investigator-assessed PFS, in line with the primary study endpoint, was extrapolated beyond the observation period in GALLIUM by an Exponential distribution, selected by investigating several alternatives modes (i.e., Log-normal, Log-logistic, Gompertz, Generalised Gamma or Weibull). This selection was based on the advice of external experts at a UK advisory board on the plausible long-term behaviour, and the observed PFS curves for patients treated with R-chemo+R in the PRIMA study (Salles et al., 2011) and the LymphoCare registry (Nastoupil et al., 2015).

PRIMA is the main Phase III, randomised controlled trial of rituximab maintenance in patients with high tumour burden FL responding to rituximab plus chemotherapy induction

(Salles et al., 2011). Roche had access to patient level data from the study and was able to analyse the outcomes for patients treated with R-chemo+R with up to 9.75 years follow up for this submission.

Post progression survival (PPS) was derived from GALLIUM for patients progressing early (early PD, up to two years from treatment initiation) and pooled between intervention and control arms, therefore assuming no difference in survival between G-chemo+G and R-chemo+R beyond progression. For patients progressing post 2 years (late PD, after two years from treatment initiation), PPS for GALLIUM was too immature, i.e. there were too few post-progression deaths. PRIMA data, where longer follow up was available, was therefore used to derive PPS for late PD.

Overall survival was an output of the Markov model as the total time in PFS and the PD states. Due to the indolent disease setting, overall survival from GALLIUM was too immature to extrapolate directly over the required time horizon.

Utility and resource use

GALLIUM utility values were available to inform utilities in PFS, presenting a very large sample of patients with previously untreated FL, compared to other studies identified in the literature. However, utilities were not collected beyond the point of progression in the study. Literature values were therefore used to inform utility in the PD states. The cross-sectional study by Wild D et al. (Wild D et al., 2006) was identified in a systematic review as the most appropriate literature source.

Resource use in the PFS state consisted of drug acquisition, administration, adverse events and supportive care. In progression, resources included supportive care and subsequent treatments. Unit costs were taken from NHS references costs 2015/16, the British National Formulary (British National Formulary, 2017), eMIT (Department of Health, 2016) and Personal Social Services Research Unit (Curtis, 2016).

Results

Table 3: Deterministic base case results

Technologies	Total Costs (£)	Total LYG	Total QALYs	Inc Costs (£)	Inc LYG	Inc QALYs	ICER (£/QALY)
G-chemo+G	[redacted]	13.25	9.96				
R-chemo+R	[redacted]	12.42	9.19	[redacted]	0.83	0.77	[redacted]

Values in the table are discounted and half cycle corrected

Chemo, chemotherapy (induction); G-chemo+G, Gazyvaro in combination with chemotherapy as induction followed by Gazyvaro monotherapy as maintenance for responders;); R-chemo+R, MabThera in combination with chemotherapy as induction followed by MabThera monotherapy as maintenance for responders ICER, incremental costs effectiveness ratio; LYG, life years gained; QALYs, quality adjusted life years

A 1,000 simulation probabilistic sensitivity analysis (PSA) was conducted in order to evaluate the uncertainty associated with the base-case estimate. The PSA indicates that G-chemo+G was cost-effective at a threshold of £30,000 per QALY gained in [redacted] of simulations compared to R-chemo+R and the probabilistic ICER agreed with the deterministic analysis.

Extensive deterministic sensitivity analyses were conducted. Varying the individual parameters produced ICERs that remained close to the base-case value in most cases.

In terms of clinical inputs, the ICER was sensitive to the choice of parametric distribution for PFS, with alternative plausible functions resulting in an ICER rage of [redacted]/QALY (Log-normal) to [redacted]/QALY (Weibull). Whereas the base case assumed a finite duration of treatment effect on PFS, the clinically plausible assumption that there is no finite duration of the treatment on progression in FL resulted in an ICER of [redacted]/QALY.

In terms of health state utilities and costs, the ICER was mainly sensitive to the assumptions on long term HRQoL not observed in GALLIUM: adjusting utilities for an age dependent decline in line with the general UK population that had not been observed in the baseline utilities in GALLIUM increased the ICER to [redacted]/QALY. In addition, using GALLIUM utilities observed at progression throughout the PD states, rather than literature values, increased the ICER to [redacted]. However, the ICER was not very sensitive to the costs of subsequent care and treatments in PD. Finally, due to the indolent nature of the disease, a significant amount of health benefits accrue over a longer time period. The ICER was therefore sensitive to the discount rate and using an alternative value of 1.5% (for costs and health effects) decreased the ICER to [redacted]/QALY.

Conclusion

The GALLIUM trial demonstrated clinical meaningful and statistical significant improvements, reducing the risk of death or progression by 34% in the primary analysis for G-chemo+G compared to R-chemo+R for previously untreated patients will FL. The de novo economic model predicted that this resulted in a median PFS increase of 2.75 years and mean increase in the time spend free of progression of 1.9 years (undiscounted) for G-chemo+G versus R-chemo+R. This PFS benefit translated in to an (undiscounted) overall survival gain of 1.45 years.

The results of the de novo cost effectiveness analysis of G-chemo+G show that it is both more effective (0.77 QALYs gained) and more costly ([redacted]) than R-chemo+R with an ICER of [redacted] /QALY.

2. The technology

2.1 Description of the technology

2.1.1 Give the brand name, UK approved name, the therapeutic class and a brief overview of the mechanism of action. For devices, provide details of any different versions of the same device.

Brand name: Gazyvaro (obinutuzumab)

Therapeutic class: ATC code: L01XC15

Gazyvaro (obinutuzumab), as a Type II anti-CD20 antibody, specifically targets the extracellular loop of the CD20 transmembrane antigen on the surface of non-malignant and malignant pre-B and mature B-lymphocytes, but not on haematopoietic stem cells, pro-B cells, normal plasma cells or other normal tissue (Goede et al., 2015, F. Hoffmann-La Roche Ltd, 2017).

CD20 is an important target for the treatment of B-cell malignancies such as FL as it is commonly expressed on most mature B-cell NHL cells, it is not shed or secreted and the degree of internalisation is generally minimal. Antibodies against CD20 deplete B-cells in lymphoid tissue and as a result, improve response rates, depth of remission, PFS, and OS in FL patients compared with chemotherapy alone (Mossner et al., 2010, Solimando et al., 2016).

Relative to Type I CD20 antibodies, e.g. MabThera (rituximab), Gazyvaro has demonstrated enhanced antibody-dependent cellular cytotoxicity (ADCC), antibody-dependent cellular phagocytosis (ADCP) and direct cell death while reducing complement dependent cytotoxicity (CDC) (Golay et al., 2013, Mossner et al., 2010). These results are consistent with those from in vivo FL models, where Gazyvaro has demonstrated stronger inhibition of tumour growth in comparison to MabThera (Dalle et al., 2011). Furthermore, glycoengineering of the fragment crystallisable (Fc) region of Gazyvaro has resulted in a higher affinity for FcγRIII (Receptor III for the Fc Region of Immunoglobulin G) receptors on immune effector cells such as natural killer cells, macrophages and monocytes as compared to non-glycoengineered antibodies, thereby enhancing ADCC and ADCP (Figure 2) (Mossner et al., 2010).

Figure 2: Mechanism of action of Gazyvaro

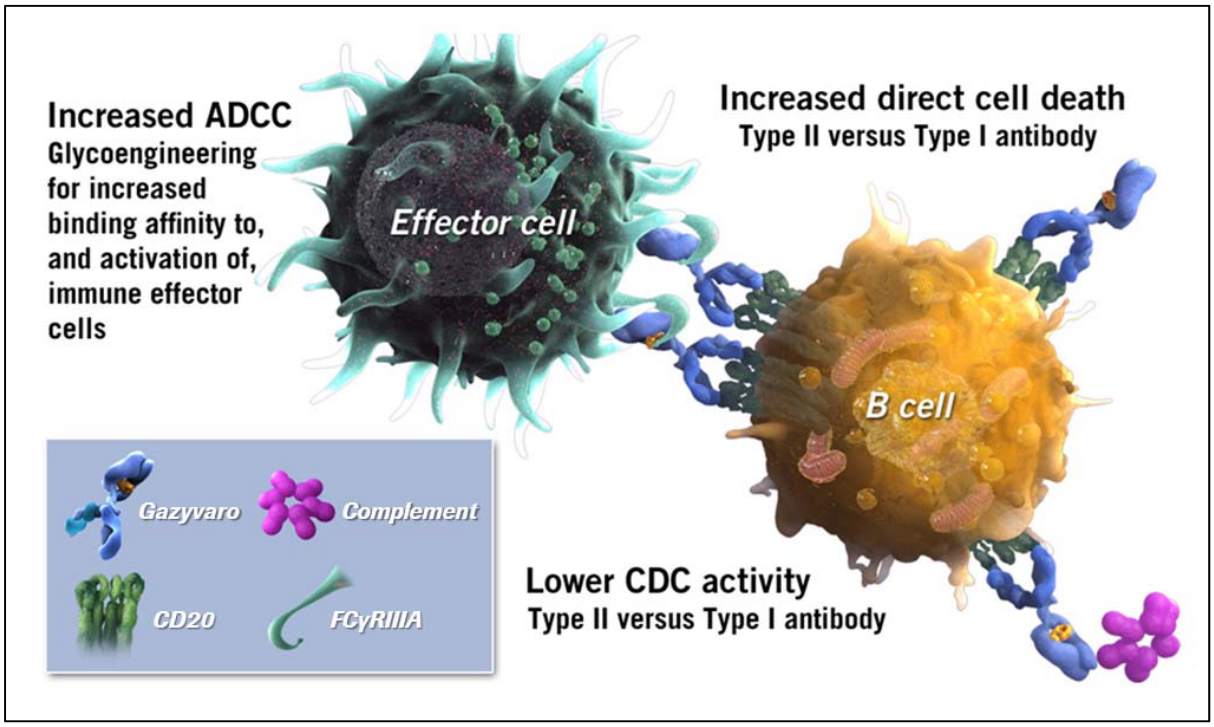


Figure adapted from (Mossner et al., 2010)
ADCC, Antibody-dependent Cellular Cytotoxicity; CDC, Complement Dependent Cytotoxicity; CD20, B-lymphocyte antigen CD20; FcγRIIIA, Receptor III for the Fc Region of Immunoglobulin G

2.2 Marketing Authorisation/CE marking and health technology assessment

2.2.1 Indicate whether the technology has a UK Marketing Authorisation/CE marking for the indications detailed in this submission. If so, give the date on which this was received. If not, state the current UK regulatory status, with relevant dates (for example, date of application and/or expected date of approval from the Committee for Human Medicinal Products).

An application for UK Marketing Authorisation was made for Gazyvaro in combination with chemotherapy, followed by Gazyvaro monotherapy as maintenance on 16th March 2017. Committee for Medicinal Products for Human Use (CHMP) opinion is anticipated in [redacted], with regulatory approval expected in [redacted].

Gazyvaro currently has Marketing Authorisation for the following indications:

- In combination with chlorambucil for the treatment of adult patients with previously untreated CLL and with comorbidities making them unsuitable for full-dose fludarabine based therapy.
- In combination with bendamustine followed by Gazyvaro maintenance for the treatment of patients with FL who did not respond or who progressed during or up to 6 months after treatment with rituximab or a rituximab-containing regimen.

Orphan designation was granted to Gazyvaro by the Committee for Orphan Medicinal Products (COMP) for CLL in October 2012, DLBCL in July 2014 and FL and marginal zone lymphoma (MZL) in June 2015.

2.2.2 Give the (anticipated) indication(s) in the UK. For devices, provide the date of (anticipated) CE marking, including the indication for use. If a submission is based on the company's proposed or anticipated Marketing Authorisation, the company must advise NICE immediately of any variation between the anticipated and the final Marketing Authorisation approved by the regulatory authorities.

The following indication wording is anticipated; however, this may be modified following comments from the CHMP:

- Gazyvaro in combination with CHOP, CVP or Bendamustine, followed by Gazyvaro maintenance therapy in patients achieving a response, is indicated for the treatment of patients with previously untreated advanced follicular lymphoma except for FL grade 3b

In the interest of breviety, the indication wording within the submission document is hereafter refererred to as Gazyvaro in combination with chemotherapy, i.e. G-chemo.

2.2.3 Summarise any (anticipated) restrictions or contraindications that are likely to be included in the (draft) summary of product characteristics (SmPC).

As noted in the draft SmPC, this medicine will be contraindicated to people who demonstrate hypersensitivity to Gazyvaro or to any of the excipients below:

- L-histidine
- L-histidine hydrochloride monohydrate
- Trehalose dihydrate
- Poloxamer 188
- Water for injections.

2.2.4 Include the (draft) SmPC for pharmaceuticals or information for use (IFU) for devices in an appendix

The draft SmPC has been included in Appendix 1.

2.2.5 Provide the (draft) assessment report produced by the regulatory authorities (that is, the European public assessment report for pharmaceuticals) and a (draft) technical manual for devices in an appendix

The current existing European public assessment for Gazyvaro (European Medicines Agency, 2016) will be updated post Marketing Authorisation to reflect the indication extension. The draft European public assessment report is provided in Appendix 2.

2.2.6 Summarise the main issues discussed by the regulatory authorities (preferably by referring to the [draft] assessment report [for example, the European public assessment report]). State any special conditions attached to the Marketing Authorisation (for example, if it is a conditional Marketing Authorisation)

The CHMP opinion has not yet been received; no special conditions will be attached to the Marketing Authorisation.

2.2.7 If the technology has not been launched, supply the anticipated date of availability in the UK.

Gazyvaro will be available in the UK for combination with chemotherapy as induction therapy followed by Gazyvaro maintenance therapy as soon as Marketing Authorisation for the proposed indication extension is received from the EMA, anticipated [redacted].

Gazyvaro is currently available in the UK for use in combination with chlorambucil for the treatment of adult patients with previously untreated CLL and with comorbidities making them unsuitable for full-dose fludarabine based therapy; and in combination with bendamustine as induction therapy followed by Gazyvaro maintenance therapy in rituximab relapsed/refractory patients with follicular lymphoma.

2.2.8 State whether the technology has regulatory approval outside the UK. If so, please provide details.

A decision as to whether an EU-wide Marketing Authorisation for the proposed indication extension in FL for Gazyvaro is anticipated in [redacted]. Gazyvaro is currently only approved for the treatment of CLL and rituximab relapsed/refractory FL in Europe and the US and does not have regulatory approval for the indication in this submission anywhere in the world.

2.2.9 State whether the technology is subject to any other health technology assessment in the UK. If so, give the timescale for completion

A Scottish Medicines Consortium appraisal is expected to begin in [redacted] for Gazyvaro in combination with chemotherapy as induction therapy followed by Gazyvaro monotherapy maintenance for previously-untreated patients with FL. Completion of this appraisal is expected no earlier than [redacted].

2.3 Administration and costs of the technology

2.3.1 For pharmaceuticals, complete the table below, indicating whether the acquisition cost is list price or includes a patient access scheme, and the anticipated care setting. For devices, provide the list price and average selling price in a similar table.

Table 4: Costs of the technology being appraised

	Cost	Source
Pharmaceutical formulation	***Gazyvaro:*** Powder for concentrate for solution for infusion	SmPC
Acquisition cost (excluding VAT)	***Gazyvaro:*** List £3,312 per 1,000 mg vial [redacted]	BNF
Method of administration	Intravenous infusion	SmPC
Doses	***Gazyvaro induction & maintenance:*** 1000 mg fixed dose •	SmPC
Dosing frequency	***Gazyvaro Induction in combination with chemotherapy (G-chemo):*** • With CHOP: 1000 mg fixed dose Gazyvaro on days 1, 8, and 15 of Cycle 1 and on Day 1 of Cycles 2–8 (21-day cycles). • With CVP: 1000 mg fixed dose Gazyvaro on days 1, 8, and 15 of Cycle 1 and on Day 1 of Cycles 2-8 (21-day cycles). • With bendamustine: 1000 mg fixed dose Gazyvaro on days 1, 8, and 15 of Cycle 1 and on Day 1 of Cycles 2–6 (28-day cycles). ***Gazyvaro maintenance:*** • 1000 mg fixed dose Gazyvaro once every 2 months for up to two years or until progression	SmPC
Average length of a course of treatment	6–8 cycles induction followed by up to 12 maintenance doses for responders to induction therapy (i.e. one maintenance dose every 2 months for up to two years or until progression)	SmPC

Average cost of a course of treatment	***Gazyvaro (List):*** • £9,936 cycle 1 • £3,312 per cycle thereafter or per maintenance dose [redacted] • [redacted]	SmPC, BNF
Anticipated average interval between courses of treatments	A person with previously untreated FL is expected to receive only one course of induction therapy followed by maintenance for responders	
Anticipated number of repeat courses of treatments	A person with previously untreated FL is expected to receive only one course of induction therapy followed by maintenance for responders	
Dose adjustments	Dose reductions are not recommended	SmPC
Anticipated care setting	Secondary care	SmPC

PAS, patient access scheme; SmPC, summary of product characteristics

2.3.2 Provide details of any patient access scheme that has been referred to NICE for inclusion in the technology appraisal by ministers and formally agreed by the company with the Department of Health before the date of evidence submission to NICE for the technology. For more information see section 5 of the NICE guide to the processes of technology appraisal.

A simple patient access scheme is in place for Gazyvaro. [redacted]

2.4 Changes in service provision and management

2.4.1 State whether additional tests or investigations are needed (for example, diagnostic tests to identify the population for whom the technology is indicated in the Marketing Authorisation) or whether there are particular administration requirements for the technology. For more information see section 5.9 of the NICE guide to the methods of technology appraisal.

There are no significant changes to the provision of services and patients management. However, Gazyvaro requires additional administration in induction in combination with chemotherapy in comparison to R-chemo. Furthermore, patients in England can be offered the subcutaneous formulation of MabThera for maintenance treatment after response to R-chemo induction, whereas Gazyvaro requires IV administration in maintenance. The respective cost implications were accounted for in the economic analysis.

2.4.2 Identify the main resource use to the NHS associated with the technology being appraised. Describe the location or setting of care (that is, primary and/ or secondary care, commissioned by NHS England specialised services and/or clinical commissioning groups), staff costs, administration costs, monitoring and tests. Provide details of data sources used to inform resource estimates and values.

Gazyvaro is administered on a 28 day cycle basis in induction with chemotherapy and every 2 months in maintenance. In induction therapy Gazyvaro is administered on days 1, 8, and 15 of cycle 1, and day 1 of cycles 2–6 (1,000 mg by intravenous infusion) (Table 5). These infusions typically take place in a hospital with an established oncology unit, which has the staffing and infrastructure required for administration of cancer treatments. Associated costs are covered by existing HRG tariffs as described in section 5.5.

Table 5: Standard infusion rate of Gazyvaro in the absence of infusion reactions/hypersensitivity

Cycle	Day of Treatment	Rate of infusion
Cycle 1	Day 1 (1,000 mg)	Administer at 50 mg/hr. The rate of infusion can be escalated in 50 mg/hr increments every 30 minutes to a maximum of 400 mg/hr.
	Day 8 (1,000 mg)	If no infusion related reaction occurred during the prior infusion when the final infusion rate was 100 mg/hr or faster, infusions can be started at a rate of 100 mg/hr and increased by 100 mg/hr increments every 30 minutes to a maximum of 400 mg/hr.
	Day 15 (1,000 mg)	
Cycles 2–6	Day 1 (1,000 mg)	
Maintenance	Every two months for two years or until disease progression (whichever occurs first)	

Reference: (F. Hoffmann-La Roche Ltd, 2017)

2.4.3 Specify if the technology requires additional infrastructure in the NHS to be put in place

No additional infrastructure is required.

2.4.4 State if and to what extent the technology will affect patient monitoring compared with established clinical practice in England

No additional monitoring is required.

2.4.5 State whether there are any concomitant therapies specified in the Marketing Authorisation or used in the key clinical trials (for example, for managing adverse reactions) administered with the technology

Premedication should be administered before Gazyvaro infusion to reduce the risk of infusion-related reactions (IRR) in patients with FL.

Table 6: Premedication to reduce risk of infusion-related reactions

Day of treatment cycle	Patients requiring premedication	Premedication	Administration
Cycle 1: Day 1	All patients	Intravenous corticosteroid* (recommended)	Completed at least 1 hour prior to Gazyvaro infusion
		Oral analgesic/anti-pyretic†	At least 30 minutes before Gazyvaro infusion
		Anti-histaminic medicine‡	
All subsequent infusions	Patients with no IRR during the previous infusion	Oral analgesic/anti-pyretic†	At least 30 minutes before Gazyvaro infusion
	Patients with an IRR (Grade 1 or 2) with the previous infusion	Oral analgesic/anti-pyretic† Anti-histaminic medicine‡	
	Patients with a Grade 3 IRR with the previous infusion OR Patients with lymphocyte counts >25 x 10^9/L prior to next treatment	Intravenous corticosteroid*	Completed at least 1 hour prior to Gazyvaro infusion
		Oral analgesic/anti-pyretic† Anti-histaminic medicine‡	At least 30 minutes before Gazyvaro infusion

Reference: (F. Hoffmann-La Roche Ltd, 2017)
*100 mg prednisone/prednisolone or 20 mg dexamethasone or 80 mg methylprednisolone.(IV administration of corticoids preferred over oral administration)
Hydrocortisone should not be used as it has not been effective in reducing rates of IRR.
†e.g. 1,000 mg acetaminophen/paracetamol
‡e.g. 50 mg diphenhydramine

2.5 Innovation

2.5.1 If you consider the technology to be innovative with potential to make a substantial impact on health related benefits that are unlikely to be included in the quality adjusted life year (QALY) calculation:

- **state whether and how the technology is a 'step change' in the management of the condition**
- **provide a rationale to support innovation, identifying and presenting the data you have used.**

Despite the indolent nature of FL, approximately 90% of newly diagnosed stage II–IV FL patients eventually require systemic treatment (Nastoupil et al., 2015). Rituximab-containing regimens are standard of care for first-line treatment of FL and maintenance therapy for two years has become accepted practice in the UK. However, approximately 20% of FL patients who receive immunochemotherapy still suffer PD within two years from diagnosis (Casulo et al., 2015b) and those patients with early progression have poorer outcomes (Maurer et al., 2016). Patients who experience early progression have a significantly increased risk of death than patients who do not relapse within the first 24 months:

- Five-year overall survival rate of 50% vs. 90% (HR 7.2, 95%: CI 4.8, 10.7), respectively, or,
- A nearly two-fold increase in the risk of death: (HR 1.89, 95% CI: 1.18, 3.03; p=0.008).

Furthermore, studies have shown that some patients with FL who progressed or relapsed after treatment with rituximab or a rituximab-containing regimen have a very short PFS of less than one year (range 5.8–10.4months) (Czuczman et al., 2012, Friedberg et al., 2008, Horning et al., 2005, Kahl et al., 2010).

Gazyvaro is a first-in-class Type II glycoengineered anti-CD20 antibody with a mode of action based on enhanced antibody dependent cellular cytotoxicity, increased direct cell death, and a lower degree of complement dependent cytotoxicity compared with non-glycoengineered, Type I antibodies such as MabThera and ofatumumab.

As highlighted in Section 4, the GALLIUM trial demonstrates that replacing MabThera with Gazyvaro in the immunochemotherapy induction and monotherapy maintenance setting for previously untreated FL patients produces a meaningful improvement in PFS. Taken together, Roche believes that Gazyvaro addresses the significant unmet need for this patient population which will provide a significant positive impact on patients' lives.

3. Health condition and position of the technology in the treatment pathway

3.1 Disease overview

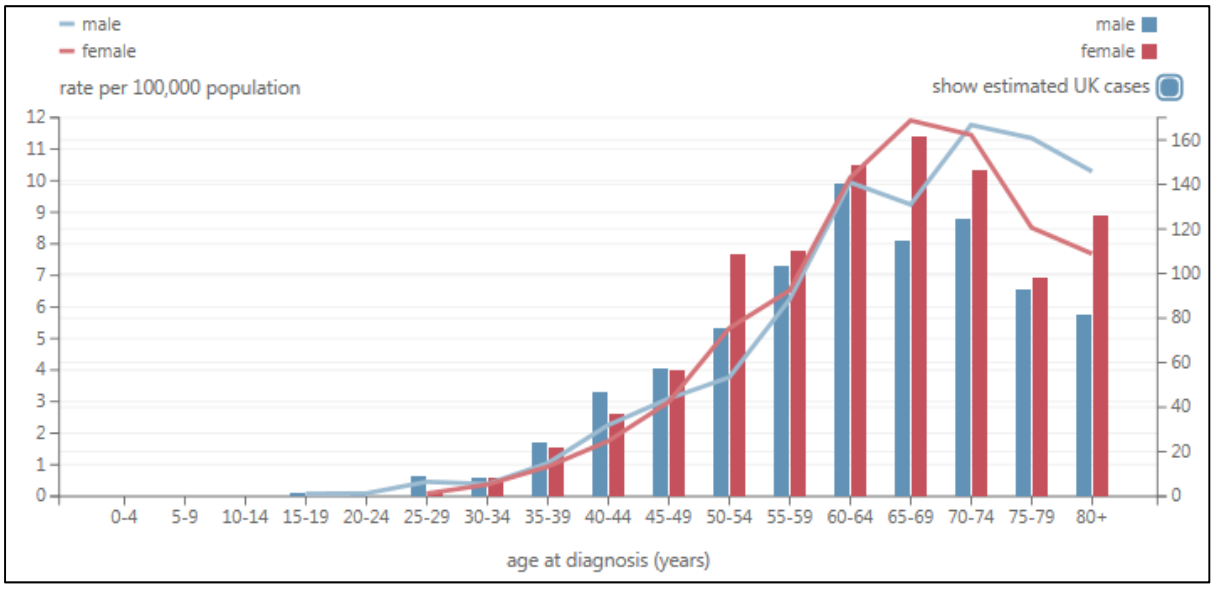
Provide a brief overview of the disease or condition for which the technology is being used. Include details of the underlying course of the disease

Non-Hodgkin lymphoma (NHL) is a heterogeneous group of lymphoproliferative malignancies, of which 80–95% of cases arise from B-cells and the remaining from T-cells (Rancea et al., 2014). NHL is divided between indolent (iNHL) and aggressive NHL subtypes (aNHL); iNHL comprises about half of all NHLs, with notable subtypes including FL, marginal zone lymphoma (MZL) and small lymphocytic lymphoma (SLL).

Incidence and prevalence of FL

FL is the second most common NHL diagnosed in Western Europe and the United States, comprising over 35% of all NHLs and 70% of indolent lymphomas (Freedman, 2015). The median age at diagnosis of FL in the UK is approximately 65 years old; the distribution of FL frequency by age (in individuals over the age of 14 years) in the UK according to the Haematological Malignancy Research Network (HMRN) database between 2004 and 2014 is presented in Figure 3 (Haematological Malignancy Research Network, 2017a).

Figure 3: Age-specific incidence of FL in the UK



The HMRN estimate that there will be 1,900 new cases of FL each year in the UK, with a higher incidence among females compared with males (Table 7). In 2015, 2,142 new cases of FL were registered in England (Office for National Statisitics, 2017). The 10-year prevalence is estimated at 15,008 cases (25.7 patients per 100,000 people), again with more females with disease than males (Table 8). This is in contrast with the prevalence estimates across all haematological neoplasms being higher in males than those of females (Haematological Malignancy Research Network, 2017c).

Table 7: Incidence of follicular lymphoma in the UK by gender

	Annual rate per 100,000[1]			M:F ratio	Median age at diagnosis	Expected UK cases per year[2]		
	Total	M	F			Total	M	F
FL	3.3	3.1	3.4	0.9	-	1900	880	1020

[1]HMRN 2004–2014
[2]Estimated by applying HMRN age and sex specific rates to 2001 UK population census strata

Table 8: Prevalence of follicular lymphoma in the UK by gender

	3-year			5-year			10-year		
	Total	M	F	Total	M	F	Total	M	F
Prevalence proportion per 100,000	10.0	9.7	10.2	15.5	14.8	16.2	25.7	24.9	26.4
Estimated UK prevalence	5822	2759	3063	9058	4205	4853	15008	7077	7931

Risk factors

In addition to clinical and environmental risk factors (e.g. medications that suppress the immune system, age, and lifestyle factors), there is increasing evidence that molecular risk factors may contribute to the risk of FL (Ma, 2012). For example, a predictive blood biomarker for FL development is the genomic translocation t(14;18)(q32;q21), which is seen in approximately 80% of patients with FL. This translocation results in the BCL-2 gene on chromosome 18 being placed on chromosome 14. A subsequent overexpression of the anti-apoptotic BCL-2 protein results in apoptosis-resistant FL cells, which therefore persist in the lymph node and eventually undergo chronic antigenic stimulation and ongoing mutagenesis processes. This may ultimately result in the proliferation of malignant clones. However, although overexpression of the BCL2 protein is detected in over 90% of cases using immunohistochemistry, it is not specific enough to ascertain a diagnosis of FL (Godon et al., 2003).

Diagnosis and staging

FL tends to be insidious in nature. Typically, patients present with multiple sites of lymphadenopathy and/or bone marrow disease [advanced-stage disease (III/IV)]. This may manifest itself with disease-related symptoms such as fatigue, weight loss, fever and night sweats (Pettengell et al., 2008). NICE Improving Outcomes Guidance for Haematological

Cancers (NG47) (National Institute for Health and Care Excellence, 2016a) provides recommendations for the integrated reporting of the diagnosis of FL, i.e. all cases should be reviewed by a haematopathologist in UK. The diagnosis of FL is typically based on the following:(Dreyling et al., 2016)

- Surgical specimen/biopsy: excisional lymph nodes are preferable; core biopsies should only be performed in the event that lymph nodes are not accessible; fine needle aspirations are inappropriate for a reliable diagnosis.
- Histological report: results and grading should be given according to the World Health Organisation (WHO) classification of FL. Grading (outlined in Table 9) is performed according to the number of centroblasts per high power field. Grades 1 through 3A are treated as indolent disease, whereas grade 3B is considered an aggressive lymphoma.
- Review: in the event of the infiltration pattern being unusual (diffuse areas, even with small cells), review by an expert haematopathologist is advised, especially for grade 3A or 3B

Table 9: Grading of follicular lymphoma

Grade	Description
1	≤5 blasts/high power field
2	6-15 blasts/high power field
3A	>15 blasts/high power field, centroblasts with intermingled centrocytes
3B	>15 blasts/high power field, pure sheets of blasts

Reference:(Dreyling et al., 2016)

FL has a characteristic immunophenotype and B-cell markers are expressed on the cell surface. The typical immunohistochemical markers for FL diagnosis (in addition to BCL-2- or CD10-) are CD20+, CD23+/-, CD10+, CD43-, BCL-2+, BCL-6+, CD5-, and CCND1-. Identification of BCL-2 gene rearrangement and translocation can facilitate diagnosis; however, the European Society for Medical Oncology (ESMO) guidelines state that biological parameters are still investigational and not yet suitable for clinical decision-making. As such, it is recommended that additional biopsy material is stored fresh frozen to allow additional molecular (currently still investigational) analyses (Dreyling et al., 2016).

Since treatment of FL heavily depends on the stage of the disease, initial staging should be thorough, particularly in the small proportion of patients presenting with early stages I and II (10%–15%) (Dreyling et al., 2016). An initial diagnostic work-up, including a physical

examination and various tests (e.g. imaging, bone marrow histology, blood tests) form the staging process. This should be documented according to the Ann-Arbor classification system (Table 10). The presence of bulky disease >5 or 6 cm may be mentioned as appropriate.

Table 10: Ann-Arbor Classification

Stage	Area of involvement
I (I_E)	One lymph node region or extralymphatic site (IE)
II (II_E)	Two or more lymph node regions or at least one lymph node region plus a single localised extralymphatic site (IIE) on the same side of the diaphragm
III (III_E, III_S)	Lymph node regions or lymphoid structures (e.g. thymus, Waldeyer's ring) on both sides of the diaphragm with optional localised extranodal site (IIIE) or spleen (IIIS)
IV	Diffuse or disseminated extralymphatic organ involvement
For all stages	
A	No symptoms
B	Unexplained fever of >38°C, drenching night swears; or loss of >10% body weight within 6 months

Reference: (Dreyling et al., 2016)

Prognosis

Patients with advanced stage FL are usually considered incurable with standard therapeutic approaches therefore treatment generally attempts to control the disease. FL is typified by a chronic course comprising of repeated relapses, treatment and progression. Generally, median life expectancy ranges have been reported from 8–12 years after diagnosis, although this has extended to around 15 years in the post-rituximab era (Tan et al., 2013).The HMRN estimate the 5-year survival rate of patients with FL in the UK to be 87.2% (Haematological Malignancy Research Network, 2017d).

Strategies to predict survival have been implemented including the Follicular Lymphoma International Predictive Index (FLIPI) (Solal-Celigny et al., 2004). A revised FLIPI, known as the FLIPI2, was developed to separate patients with significantly different hazard ratios for progression/relapse in the era of anti-CD20 monoclonal antibody treatments (Federico et al., 2009). An overview of the differences between the FLIPI and FLIPI2 indexes is outlined in Table 11.

Table 11: FLIPI and FLIPI2 prognostic indexes for follicular lymphoma

FLIPI (retrospective analysis; pre-rituximab era)			FLIPI2 (prospective analysis; rituximab era)		
5 factors	Age	<60 years vs ≥60 years	**5 factors**	Age	<60 years vs ≥60 years
	Haemoglobin	≥12g/dL vs <12g/dL		Haemoglobin	≥12g/dL vs <12g/dL
	Serum LDH	≤ULN vs >ULN		Serum β-2 microglobulin	≤ULN vs >ULN
	Ann-Arbor stage	I-II vs III-IV		Bone marrow involvement	absent vs present
	No. of nodal sites	≤4 vs >4		Longest diameter of largest lymph node	≤6 cm vs >6 cm

Risk group	No. of FLIPI factors	5-yr OS (%)	10-yr OS (%)	Relative risk	Risk group	No. of FLIPI2 factors	3-yr PFS (%)	3-yr OS (%)	5-yr PFS (%)
Good	0–1	91	71	1	Good	0	91	99	79
Intermediate	2	78	51	2.3	Intermediate	1–2	69	96	51
Poor	≥3	53	36	4.3	Poor	≥3	51	84	20

FLIPI, Follicular Lymphoma International Predictive Index; LDH, lactate dehydrogenase; OS, overall survival; PFS, progression-free survival; ULN, upper limit of normal

The FLIPI2 is considered to be predictive of treatment outcomes in patients with newly diagnosed FL who receive immunochemotherapy; however, the scores do not provide guidance or assist in predicting the immediacy or type of the treatment that should be given in the FL setting (Bello et al., 2012).

Early identification of FL patients at high risk of relapse is important for treatment optimisation. To this end, detection of minimal residual disease (MRD) has emerged as a potentially important tool for the detection of persistent residual tumour cells, and the evaluation of treatment effectiveness and long-term prognosis in this patient population (Lobetti-Bodoni et al., 2013). MRD can be detected using different methods (Faham et al., 2012); currently most of the available information on FL derives from polymerase chain reaction (PCR)-based MRD detection (Lobetti-Bodoni et al., 2013). MRD analysis by PCR investigates the persistence of residual tumour cells through the amplification of a tumour-specific molecular marker, for which in the case of FL the t(14;18) translocation is particularly suitable.

Studies have demonstrated the predictive value of MRD detection in FL; for example, patients treated with rituximab in addition to CHOP were shown to have a better clinical outcome (defined as freedom from recurrence) if they achieved MRD negative status compared with those who never achieved or lost molecular remission (57% vs 20%, p<0.001) (Rambaldi et al., 2002). The prognostic role of MRD is under exploration; MRD negativity predicted better PFS in patients with FL receiving rituximab maintenance after induction chemotherapy, while MRD positivity at the end-of-induction was an independent adverse predictor (Ladetto et al., 2013). Furthermore, in a study of rituximab-refractory FL patients, MRD negativity has been shown to prognostically identify a group of patients who appear to benefit from treatment with immunochemotherapy at relapse, and moreover, patients without MRD negativity at end-of-induction appear to have a poorer prognosis (Pott et al., 2015).

3.2 Effects of the disease on patients, carers and society

Disease burden

Indolent NHL is usually considered incurable with standard therapeutic approaches therefore treatment approaches focus on attempting to control the disease for the longest period while maintaining quality of life. FL commonly presents as painless, slow progressing adenopathy. While systemic symptoms such as fever, night sweats, or weight loss >10% are infrequent early in the disease, they can be observed at later stages. Similarly, symptoms related to bone marrow dysfunction, such as anaemia, leukopenia, or thrombocytopenia, may also be observed in the later stages of the disease (Medscape., 2016). At diagnosis, the majority of people with FL have advanced (stage III-IV Ann Arbor stage disease); bone marrow involvement is also common and present in more than 50% of patients.

The progression of FL varies among patients depending on the speed of tumour growth and involvement of other organs. Approximately 20% of FL patients who receive immunochemotherapy still suffer PD within two years from diagnosis (Casulo et al., 2015b). Patients who experience early progression have a significantly increased risk of death than patients who do not relapse within the first 24 months:

- Five-year OS rate of 50% vs. 90% (HR 7.2, 95% CI 4.8, 10.7), respectively, or
- A nearly two-fold increase in the risk of death: HR 1.89 (95% CI 1.18, 3.03; p=0.008)

These findings were further confirmed with a validation set (Casulo et al., 2015b) and a subsequent smaller independent study (n=94) of FL patients with clinical stage II–IV who

had received immunochemotherapy (Murakami, Kato et al. 2016). The latter estimated that the five-year OS of patients with and without early PD was 48% (95% CI: 21, 71) and 96% (95% CI: 88, 99), respectively (P<0.0001). In a multivariate Cox regression analysis, early PD remained a significant independent prognostic factor (HR 7.82 for death, 95% CI: 1.97, 31.0, p=0.003), even when adjusting for previously validated prognostic factors such as FLIPI (early PD HR 11.2, 95% CI: 3.13, 40.3, p<0.001) or FLIPI2 (early PD HR 13.5, 95% CI 3.22, 56.3, p<0.003) and high tumour burden (Murakami et al., 2016).

Furthermore, some cases of FL will transform to more aggressive forms such as DLBCL; recent studies report a risk of transformation of about 2% to 3% per year through at least 10 to 15 years of diagnosis (Casulo et al., 2015a). Transformation has been shown to severely worsen outcome with 10-year survival decreasing to 36% for patients with transformed FL (Al-Tourah AJ et al. 2008). However, recent studies suggest this may be improving with the use of rituximab; median overall survival for all patients after transformation was 50 months, and at 5 years, overall survival was 73% in patients treated with R-CHOP (Link et al., 2013).

Several clinical factors have been found to be associated with a higher risk of future transformation, including non-response to first-line FL therapy (HR 2.5, 95% CI: 1.5, 4.2), increased lactate dehydrogenase (HR 2.5; p=0.0013), high FLIPI score (HR 2.1, 95% CI: 1.3, 3.4; p=0.002) and advanced disease stage (p=0.002) (Alonso* et al., 2015, Al-Tourah et al., 2008, Link et al., 2013). In the PRIMA study, after a median six-year follow-up, patients with histologic transformation had less frequent CR (50.3% vs. 67.4%; p=0.03) and more disease progression (28.2% vs. 9.6%; p<0.001) than patients without histologic transformation. The estimated median OS for patients with histologic transformation was shorter than for patients without histological transformation (3.8 vs. 6.4 years; HR 3.9, 95% CI: 2.2, 6.9) (Sarkozy et al., 2016).

Patient quality of life

Patients with FL also experience chronic disease pathology with repeated relapses which are often unpredictable and require repeated courses of treatment (Blaes A et al., 2011). The toxicity of treatment and symptoms related to repeated relapse in iNHL may have substantial burden on the health-related quality of life (HRQoL) of patients (Cheung et al., 2009).

When comparing iNHL with aNHL, studies have reported no significant difference in HRQoL between the patient groups. This suggests that although indolent in nature, iNHL has a similar impact on a patient's HRQoL as the more aggressive forms of NHL. A study using the

short-form health survey (SF-36) and the FACT-F PRO instruments found that there was, overall, no significant difference between the two patient groups (Blaes A et al., 2011). An exception to this was SF-36 results showing that aNHL patients were significantly more impaired in physical function (specifically addressing physical limitations) than those with iNHL. Although it is not clear whether complications from chemotherapy contributed to this, the lower level of active therapy in the iNHL arm may explain why their physical function appeared significantly better.

In a study of 222 FL patients, those who had relapsed were more likely to experience a reduction in HRQoL than other patients (Pettengell et al., 2008). Patients were divided into those receiving chemotherapy and those who were not and were analysed for HRQoL across five disease states: 'newly diagnosed', 'active disease–relapsed', 'partial response', 'complete response' and 'disease free'. Based on results of four of the five patient-reported outcome (PRO) measures, participants receiving treatment reported overall, statistically significant worse health functioning (p=0.004), depressive symptoms (p=0.005) and activity impairment (p=0.009) than those participants in remission.

The results of this study demonstrate the importance of remission in terms of HRQoL. Patients who have relapsed are more likely to experience worse HRQoL and other patient-reported health outcomes than patients newly diagnosed, in partial or complete remission or when completely disease free, thereby demonstrating the importance to achieve and maintain disease control.

Impact on carers

Further to the impact on patients, iNHL places a significant burden on their families and friends. A cross-sectional survey of iNHL patients identified that almost one-quarter of patients depended on caregiver assistance, with the majority (74%) being unpaid care provided by a spouse, partner, relative or friend. Caregivers face an increasing burden of physical, psychological and family life disruptions, in addition to the economic burden resulting from reduced time at work (Cheung et al., 2009).

Impact on society

In addition to the direct costs of therapy for iNHL, the indirect costs associated with the impact on the workforce and burden to caregivers are borne by the society. Studies have demonstrated that iNHL may result in early retirement, while patients who remain in work may suffer from reduced productivity due to symptoms related to disease progression, side

effects of treatment, hospitalisation for complications, time taken for follow-up visits and tests, and emotional stress. The impact on productivity and ability to participate in daily activities can be variable at different stages of disease; however patients receiving active chemotherapy were noted to have considerable impairment in these aspects compared with those patients in the initial observation period of therapy or patients who reached a first remission or subsequent remission (Cheung et al., 2009).

3.3 Clinical pathway of care

Present the clinical pathway of care that shows the context of the proposed use of the technology. This information may be presented in a diagram. Explain how the new technology may change the existing pathway. If a relevant NICE clinical guideline has been published, the response to this point should be consistent with the guideline and any differences should be explained.

NICE clinical guidelines for the diagnosis and management of NHL (National Institute for Health and Care Excellence, 2016d) suggests the following treatment recommendations for FL (outlined in Table 12 with the management of first-line FL summarised in Figure 4).

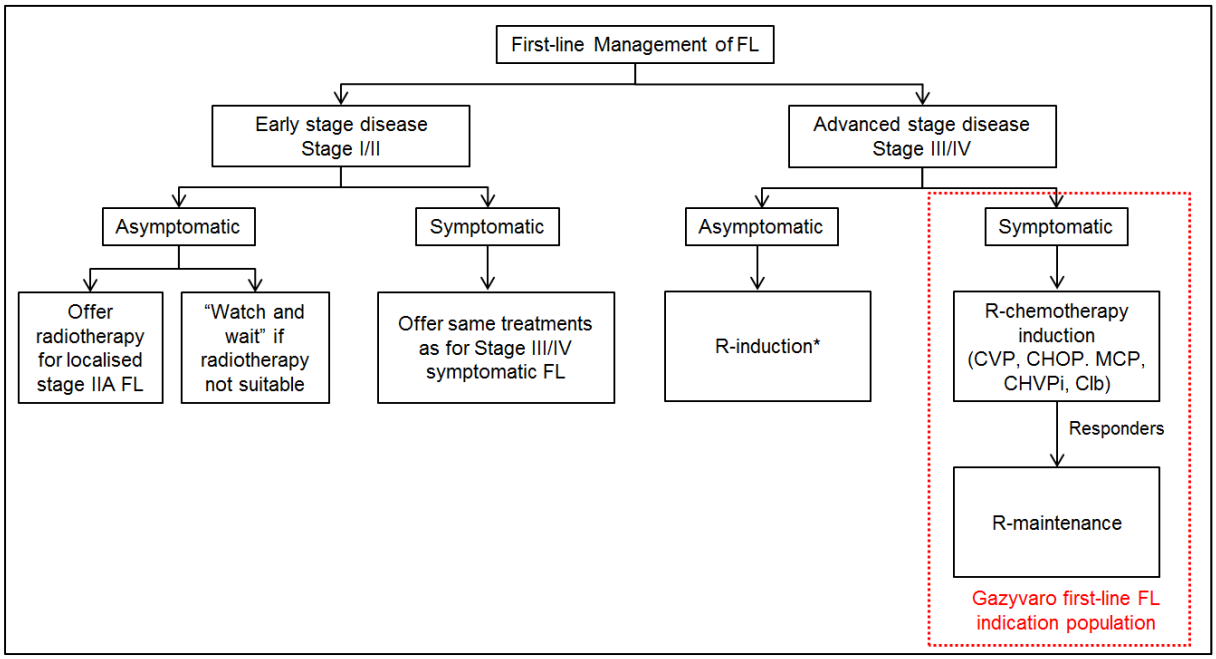
Table 12: NICE guidelines for the management of follicular lymphoma (NG52)

Stage of disease	Treatment/management recommendation
First-line treatment for stage IIA FL	• Local radiotherapy in people with localised stage IIA FL • Consider "watch and wait" for people with stage IIA FL who are asymptomatic and for whom treatment with a single radiotherapy volume is not suitable • Offer the same treatments that might be offered to people with advanced-stage (stages III and IV) symptomatic FL to people with stage IIA FL who are symptomatic and for whom radiotherapy is not suitable
Advanced-stage asymptomatic FL	• Offer rituximab induction therapy to people with stage III or IV FL who are asymptomatic
Advanced-stage symptomatic FL ***[Position of proposed new technology]***	• Rituximab in combination with chemotherapy (CVP, CHOP, MCP, or chlorambucil) for the treatment of symptomatic stage III and IV FL in previously untreated people (NICE TA243) • Rituximab maintenance is recommended as an option for the treatment of people with FL that has responded to first-line induction therapy with rituximab in combination with chemotherapy (NICE TA226)
Advanced-stage relapsed or refractory FL (NICE TA137)	• Rituximab, within its Marketing Authorisation, in combination with chemotherapy, is recommended as an option for the induction of remission in people with relapsed stage III or IV FL

	• Rituximab monotherapy as maintenance therapy, within its Marketing Authorisation, is recommended as an option for the treatment of people with relapsed stage III or IV FL in remission induced with chemotherapy with or without rituximab • Rituximab monotherapy, within its Marketing Authorisation, is recommended as an option for the treatment of people with relapsed or refractory stage III or IV FL, when all alternative treatment options have been exhausted (that is, if there is resistance to or intolerance of chemotherapy)
Consolidation therapy	• Offer consolidation with autologous stem cell transplantation for people with FL in second or subsequent remission (complete or partial) who have not already had a transplant and who are fit enough for transplantation • Consider consolidation with allogeneic stem cell transplantation for people with FL in second or subsequent remission (complete or partial): o who are fit enough for transplantation and o for whom a suitable donor can be found and o when autologous stem cell transplantation has not resulted in remission or is inappropriate (for example, because stem cell harvesting is not possible).
Transformed FL	• Consider consolidation with autologous stem cell transplantation for people with transformation of previously FL that has responded to treatment and who are fit enough for transplantation • Consider consolidation with autologous or allogeneic stem cell transplantation for people with transformation of follicular lymphoma who need more than 1 line of treatment for a response and who are fit enough for transplantation • Do not offer consolidation with high-dose therapy and autologous or allogeneic stem cell transplantation to people presenting with concurrent diagnoses of FL and DLBCL that have responded to first-line treatment

ASCT, autologous stem cell transplantation; DLBCL, diffuse large B-cell lymphoma; FL, follicular lymphoma

Figure 4: NICE algorithm for the first-line management of FL



*rituximab does not have a UK Marketing Authorisation for this indication

First-line treatment options for advanced-stage, symptomatic FL

Rituximab induction with chemotherapy

In patients with advanced stage symptomatic FL, rituximab, in combination with chemotherapies has transformed the course of treatment, delivering major improvements in survival with a manageable toxicity profile. Rituximab has become a universal standard of care for the treatment of FL (Fisher et al., 2005). Improvements in both PFS and OS with R-chemo have been demonstrated in multiple randomised clinical trials, compared with chemotherapy alone (Herold et al., 2007, Hiddemann et al., 2005, Hochster et al., 2009, Marcus et al., 2005, Marcus et al., 2008, Salles et al., 2008).

Rituximab induction plus monotherapy maintenance

The benefit of maintenance rituximab after induction chemotherapy, in comparison with observation, has been demonstrated in the Phase III PRIMA trial with advanced FL patients (Salles et al., 2011, Seymour JF et al., 2013). Patients receiving rituximab maintenance, following R-chemo, had a six-year PFS rate of 59% compared with 43% in the observation group (HR 0.58, 95% CI: 0.48, 0.69 p<0.0001). There was no difference in OS between study arms: 11.3% of patients had died in the observation arm (6-year OS estimate 88.7%) compared to 11.7% in the rituximab maintenance group (6-year OS estimate 87.4%) (Seymour JF et al., 2013). The rate of AEs was low, and haematologic toxicity induced

during chemotherapy treatment improved in both rituximab maintenance and observation groups. Overall, the study showed that rituximab maintenance did not negatively impact quality of life or treatment-related symptoms when compared to the observation group, and in both groups the toxicity induced during chemotherapy treatment at induction improved after discontinuation of chemotherapy post-induction (Zhou et al., 2014).

A systematic literature review with a meta-analysis, assessing several trials in which a rituximab maintenance regimen was compared to observation, following induction in the first-line setting, confirmed a clear PFS benefit for the rituximab maintenance strategy with a pooled HR of 0.54 (95% CI: 0.48, 0.60). Although maintenance rituximab after induction chemotherapy has yet to demonstrate an OS benefit in randomised clinical trials(HR 0.86, 95% CI: 0.60, 1.25) (Vidal et al., 2011), real-world data from a Danish population-based cohort has shown that patients consolidated with rituximab maintenance significantly improved 5-year OS compared to patients not receiving maintenance (90% vs 83%, p=0.003) (Madsen C et al., 2016).

Regimens used in the UK 2014

The regimens used in the UK for the treatment of indolent NHL as generated by the Systemic Anti-Cancer Therapy (SACT) chemotherapy dataset for 2014 are summarised below. The major caveat to interpreting these data is that they are not specific to FL and include all lines of treatment; therefore, they cannot be used to confirm which regimens are specifically used for the first-line treatment of advanced FL

Table 13: SACT dataset for indolent NHL treatment regimens used in the UK, 2014

Regimen	Total patients	First cycles	Total cycles
MabThera (rituximab)	1,508	843	5,085
R-CVP	441	327	1,881
R-bendamustine	420	352	1,554
R-CHOP	305	244	1,188
Bendamustine	86	66	355
GALLIUM trial	53	10	263
Other trials	51	29	226
Chlorambucil	69	51	221
Cladribine	103	99	186
R-chlorambucil	49	38	181

Market research to determine the immunochemotherapy regimens used as first-line treatment for FL in UK clinical practice has revealed the following (Roche Products Ltd.):

Table 14: FL first-line regimens in UK clinical practice, Q1 2017

Regimen	Proportion use, %
Induction	
R-CVP	36
R-CHOP	13
R-bendamustine	29
R-FC	8
R-other	2
FC	11
Other	1

This information has been corroborated at an Advisory Board by external experts[3], who informed Roche that R-CVP and R-bendamustine are the most commonly used induction immunochemotherapy regimens in the UK, with R-CHOP retained for use in patients at high-risk of transformation.

The unmet need in FL

Advanced FL is an indolent orphan disease that is considered incurable. Certain patient subgroups have significantly poorer outcomes, such as those who suffer early PD or transformation following diagnosis or treatment initiation; high-risk FLIPI/FLIPI2 groups and patients with advanced age. Despite the indolent nature of FL, approximately 90% of newly diagnosed stage II-IV FL patients eventually require systemic treatment (Nastoupil et al., 2015), and approximately one fifth of FL patients receiving immunochemotherapy still suffer a PD event within two years. These patients have a significantly increased risk of death/significantly worse OS than patients who do not relapse within the first 24 months (Casulo et al., 2015b, Murakami et al., 2016). Furthermore, FL patients who suffer histological transformation into aggressive lymphoma within 18 months of diagnosis have significantly poorer OS than patients who transform later (22% vs. 66%) and those patients who never transform (Al-Tourah et al., 2008, Schatz et al., 2013, Wagner-Johnston et al., 2015).

R-chemo is the current standard of care in first-line FL patients, yet approximately 50% progress or die within six years of treatment initiation with rituximab-based therapy (Salles, 2016, Salles et al., 2011, Tan et al., 2013). As FL is an incurable disease, patients suffer successive relapses. Even though therapies are available in second and later lines, responses to treatment become progressively shorter with each subsequent relapse.

[3] An expert advisory board was consulted at a one-day meeting in April 2017. The panel consisted of consultant haematologists specialising in the management of patients with FL, many of whom have experience of Gazyvaro from clinical trials. The panel was selected based on their significant clinical and research experience.

Moreover, successive treatments are associated with increased risk of cumulative toxicity and secondary malignancies. Therefore, there is a need for first-line FL treatments that can result in longer remissions and longer time to next lymphoma treatment, and fewer patients requiring treatment in a relapse setting.

Gazyvaro is a Type-II anti-CD20 antibody that has demonstrated enhanced antibody-dependent cellular cytotoxicity, antibody-dependent cellular phagocytosis and direct cell death while reducing complement dependent cytotoxicity, relative to Type-I antibodies (i.e. rituximab). Evidence for how Gazyvaro can address this unmet need in previously untreated patients with advanced FL was generated in the Phase III open-label GALLIUM (BO21223) study and is discussed in detail in Section 4 of this submission.

3.4 Life expectancy of people with the disease in England

Provide information about the life expectancy of people with the disease or condition in England Include the source of the data. Please provide information on the number of people with the particular therapeutic indication for which the technology is being appraised. If the Marketing Authorisation also includes other therapeutic indications for the technology, provide information about the numbers of people with these diseases or conditions in England and provide the source of the data. This is to assess whether the technology may be suitable for consideration as a 'life-extending treatment at the end of life' as described in section 6.2.10 of the NICE guide to the methods of technology appraisal.

People with FL may live for many years after diagnosis. In England and Wales, Cancer Research UK notes that approximately 90% of stage I and stage II patients with FL survive for 5 years or more following diagnosis. The 5-year survival rate declines to approximately 80% in patients with stage III or IV disease (Cancer Research UK, 2017).

These statistics are corroborated by the HMRN, which cite 5-year survival rates (based on data accrued from 2004–2014) of 87.2%. The HMRN report no difference in survival rates between men and women, although 5-year survival decreases with age, particularly in patients aged 80 years and older (50–55%) (Haematological Malignancy Research Network, 2017d).

Data from the US National LymphoCare study has demonstrated that patients who progress within 2 years of treatment have a poorer prognosis than responders; 5-year survival among patients with progressive disease was 50% compared with 90% for patients responding to a rituximab-containing regimen (Casulo et al., 2015b).

3.5 Guidance related to the condition

Provide details of any relevant NICE guidance, pathways or commissioning guides related to the condition for which the technology is being used. Specify whether any subgroups were explicitly addressed.

Details of relevant NICE guidance for the diagnosis and management of previously-untreated FL are listed below, based on the final scope.

- Haematological cancers: improving outcomes [NG47] (National Institute for Health and Care Excellence, 2016a)
- Non-Hodgkin's lymphoma: diagnosis and management [NG52] (National Institute for Health and Care Excellence, 2016d)
- Rituximab for the first-line treatment of stage III-IV follicular lymphoma [TA243] (National Institute for Health and Care Excellence, 2012)
- Rituximab for the first-line maintenance treatment of follicular non-Hodgkin's lymphoma [TA226] (National Institute for Health and Care Excellence, 2011)

3.6 Other clinical guidelines

Provide details of other clinical guidelines (for example, UK guidance from the royal societies or European guidance) and national policies.

British Committee for Standards in Haematology (BCSH) Guidelines on the Investigation and Management of FL (UK) (2012)

The BCSH provides up-to-date evidence-based guidelines for both clinical and laboratory haematologists for the diagnosis and treatment of haematological disease. The guidelines are written according to the BCSH process by a team of expert consultants and clinical scientists currently practicing in the UK. Below are the BCSH recommended treatments by disease stage (McNamara et al., 2012).

Early Stage Disease

Exclude the more advanced disease and record the FLIPI index. Early stage FL comprising Ann Arbor stage I and stage II disease, where the involved nodes are contiguous, should be treated with local radiotherapy. The recommended current standard dose is 24 Gy in 12 fractions.

Combined modality treatments (radiotherapy plus chemotherapy) are recommended as an alternative approach, since the majority of relapses occur outside the radiation field, and are seen in up to 50% of patients. Patients with limited stage FL should be observed without treatment, especially if there are concerns with regards to side effects from involved field radiotherapy (e.g. fertility preservation in young women, elderly frail patients where there is significant morbidity) (McNamara et al., 2012).

Advanced Stage

A "watch and wait" approach should be limited to asymptomatic advanced stage FL (Ann Arbor stage III / IV or stage IIB) in an attempt to delay the need for chemotherapy.

R-chemo should be used in patients with newly diagnosed, symptomatic advanced stage FL who require therapy. There is no strong evidence to support one regimen over another. Rituximab maintenance, after response to induction therapy, prolongs PFS and is recommended in patients responding to induction rituximab-based chemotherapy.

Autologous stem cell transplantation (autoSCT) does not have a role in first-line therapy for FL outside a clinical trial and R is not recommended (McNamara et al., 2012).

ESMO Clinical Practice Guidelines (Europe)

The most recent European guidelines recommend that the therapeutic approach in FL should be decided based upon the assessment of clinical risk factors (such as FLIPI/FLIPI2), symptom burden and patient perspective (Dreyling et al., 2016).

Table 15: ESMO clinical practice guidelines for the treatment of FL

Low Tumour Burden		High Tumour Burden*	
Stage I/II	**Stage III/IV**	**Stage III/IV (<65 years†)**	**Stage III/IV (>65 years)**
Front Line			
Radiotherapy (involved field) 24-36 Gy In selected cases, "watch and wait" or rituximab monotherapy	"Watch and wait" In selected cases, consider rituximab monotherapy	Chemoimmunotherapy (e.g. R-CHOP, R-CVP, R-benda) In selected cases, rituximab monotherapy CR/PR Rituximab maintenance (every 2 months, up to 2 years)	Chemoimmunotherapy (e.g. R-CVP, R-benda, R-CHOP or brief chemoimmunotherapy In selected cases, rituximab-chlorambucil rituximab monotherapy CR/PR Rituximab maintenance (every 2 months, up to 2 years)

Relapse/progress			
"Watch and wait" Rituximab monotherapy In selected cases, palliative radiation (e.g. 2 x 2 Gy)	"Watch and wait" Rituximab monotherapy Chemoimmunotherapy (e.g. B-Benda, R-CHOP, R-CVP) Idelalisib (double refractory cases)	Dependent on first-line regimen and remission duration Chemoimmunotherapy (e.g. R-benda, R-CHOP, R-CVP) Discuss high-dose consolidation with ASCT Rituximab maintenance (every 3 months, up to 2 years) RIT or rituximab monotherapy in selected cases discuss allogeneic transplantation Idelalisib (double refractory cases)	Dependent on first-line regimen and remission duration Chemoimmunotherapy (e.g. R-benda, R-CHOP, R-CVP) Rituximab maintenance (every 3 months, up to 2 years) RIT or rituximab monotherapy Idelalisib (double refractory cases)

R-CHOP, rituximab, cyclophosphamide, doxorubincin, vincristine, prednisolone; R-CVP, cyclophosphamide, vincristine and prednisolone; R-benda, rituximab plus bendamustine; CR, complete response; ASCT, autologous stem cell transplantation.
*High tumour burden is defined as the presence of one of the following: bulk (>7cm) disease or three affected lymph nodes in distinct areas of >3cm; symptomatic splenic enlargement; organ compression; elevated LDH or β2M, or B symptoms.
†According to biological age

3.7 Issues relating to current clinical practice

Describe any issues relating to current clinical practice, including any variations or uncertainty about established practice.

Despite multiple therapeutic options being available, no curative treatment exists for the majority of patients with FL. In the UK, all treatment decisions are individualised and made within an Improving Outcomes Guidance (IOG) compliant haemato-oncology multi-disciplinary team setting (National Institute for Health and Care Excellence, 2016a).

Although radiotherapy is recommended as first-line therapy for patients with a low tumour burden, adherence to the standard of care is low (Friedberg et al., 2012). The uncertainty concerning the most adequate therapy for early-stage FL may in part result from the observation that a watch and wait approach may not be inferior to radiotherapy in these patients (Advani et al., 2004). Clinical experts have confirmed that all symptomatic patients with advanced FL (stage III–IV) require treatment and receive immunochemotherapy followed by rituximab maintenance

Several advances in identifying prognostic markers have been made in recent years. However, while the FLIPI-2 prognostic index was found to be predictive of treatment outcomes in newly diagnosed FL patients who received immunochemotherapy, it does not provide sufficient insights on when to initiate treatment and which therapy to use (Bello et al., 2012). Modern cell biologic and genetic techniques including gene sequencing provide a deeper understanding of the pathophysiology of FL and may facilitate the identification of distinct biologic subgroups, but current data are inconsistent and cannot be applied for treatment guidance in clinical practice today (Hiddemann and Cheson, 2014).

While initial therapy has shown to be highly efficacious, approximately 20% of FL patients who receive immunochemotherapy still suffer PD within two years from diagnosis; these patients have been shown to have a poorer prognosis (Casulo et al., 2015b, Maurer et al., 2016). The US National LymphoCare study (analysing 588 patients with stage 2–4 FL having received first-line R-CHOP) demonstrated that 5-year survival among patients with disease progression within 2 years of treatment was lower compared with those without disease progression, 50% vs 90% respectively (Casulo et al., 2015b).

Furthermore, studies have shown that some patients with FL who progressed or relapsed after treatment with rituximab or a rituximab-containing regimen have a very short PFS of less than one year (range 5.8–10.4months) (Czuczman et al., 2012, Friedberg et al., 2008, Horning et al., 2005, Kahl et al., 2010).

Moreover, the majority of patients with advanced-stage FL will experience disease progression after a period of several years. This remains the case despite the benefit of additional rituximab in the form of maintenance therapy, as demonstrated in the PRIMA study where approximately 41% of patients had disease progression after 6 years follow-up (Seymour JF et al., 2013).

Similar data was reported by Tarella C et al. in which data recorded in an Italian registry from 574 patients with low grade NHL (diagnosed and managed since 2000) who had undergone

primary treatment were analysed. Five-year overall survival projections for responsive patients were 83.5% compared with 60.7% for patients with early disease progression (within 6 months of treatment) and 48.4% for fully refractory patients (stable or progressive disease following frontline therapy that was either completed or discontinued in order to intensify treatment) (Tarella et al., 2014).

Given the poor prognosis for patients who progress early and the consistent frequency of relapses to rituximab-containing regimens, there is a need for alternate first-line therapies.

3.8 Equality Issues

Provide an assessment of whether the use of this technology is likely to raise any equality issues. Please document if there are any potential issues that: could exclude from full consideration any people protected by the equality legislation who fall within the patient population for whom the technology is or will be licensed could lead to recommendations that have a different impact on people protected by the equality legislation compared with the wider population, for example by making it more difficult in practice for a specific group to access the technology could lead to recommendations that have any adverse impact on people with a particular disability or disabilities. Please provide any evidence that would enable the Committee to identify and consider the impact of equality issues. State how the analysis has addressed these issues.

No equality issues have been identified.

4. Clinical effectiveness

4.1 Identification and selection of relevant studies

4.1.1 Methodology and objective. Advise whether a search strategy was developed to identify relevant studies for the technology. If a search strategy was developed and a literature search carried out, provide details under the subheadings listed in this section. Key aspects of study selection can be found in Systematic reviews: CRD's guidance for undertaking reviews in health care (University of York Centre for Reviews and Dissemination).

A systematic literature review (SLR) was conducted to identify all relevant published and unpublished RCT evidence relating to relating to the use of Gazyvaro in previously-untreated FL.

The SLR was conducted according to the NICE guide to the methods of technology appraisal 2013 and therefore adhered to the Centre for Reviews and Dissemination guidance for undertaking systematic reviews in health care.

The systematic search was run on electronic databases (i.e. MEDLINE, MEDLINE In-Process, EMBASE and Cochrane) and was supplemented by hand searches to ensure that all relevant studies had been included.

The search for controlled trials was limited to capture publications from 1998 onwards as this coincided with the market approval date of MabThera.

MEDLINE (including MEDLINE In-Process), EMBASE, and Cochrane searches were conducted on 23rd June 2016 and updated on 6th March 2017. Each database was searched individually.

4.1.2 Describe the search strategies used to retrieve relevant clinical data. The methods used should be justified with reference to the decision problem. Sufficient detail should be provided so that the results may be reproduced. This includes a full list of all information sources and the full electronic search strategies for all databases, including any limits applied. The search strategies should be provided in an appendix

The complete search strategy for this review is provided in Appendix 3. The following sources were searched, using search terms that combined population, interventions and study types:

- Electronic databases, searched separately from 1998 onwards:
    - EMBASE
    - MEDLINE and MEDLINE In-Process
    - Cochrane Central Library of Controlled Trials (Cochrane Library)
- Congress proceedings were also searched manually from 2014 onwards:
    - American Society of Hematology (ASH) Annual Meeting
    - American Society of Clinical Oncology (ASCO)
    - European Society for Medical Oncology (ESMO)
    - European Haematology Association (EHA)
    - International Conference on Malignant Lymphoma
- Clinical trial registries were also searched (not restricted by time period):
    - ClinicalTrials.gov of the US National Institute of Health (NIH)
    - WHO's meta-registry "International Clinical Trials Registry Platform Search Portal" (ICTRP)
    - EU Clinical Trials Register
    - Klinische Prüfungen PharmNet
- Cancer association networks
    - National Comprehensive Cancer Network (NCCN)
    - European Organisation for the Research and Treatment of Cancer (EORTC)
    - ESMO International Network of Agencies for Health Technology Assessment
- HTA agency websites:
    - International Network of Agencies for Health Technology Assessment (INAHTA)
    - National Institute for Health and Care Excellence (NICE)
    - National Institute for Health Research HTA
    - Canada: Canadian Agency for Drugs and Technologies in Health (CADTH)

4.1.3 Study selection. Describe the inclusion and exclusion selection criteria, language restrictions and the study selection process in a table. Justification should be provided to ensure that the rationale for study selection is transparent.

4.1.3.1 Inclusion and exclusion criteria

The eligibility criteria used for the systematic review are presented in Table 16 below; no language restrictions were used.

Table 16: Eligibility criteria for systematic literature review of RCT evidence

Domain	Inclusion criteria	Exclusion criteria
Population	Patients with previously untreated iNHL	• Not focussing on human data • Not iNHL • Not previously untreated iNHL
Interventions and comparators	All licensed and investigative interventions	Not including treatment of interest
Outcomes	All primary and secondary outcomes available, including all efficacy, all end-points, PROs, HRQoL outcomes, and safety Examples include but are not restricted to: • Efficacy endpoints reported in studies, including PFS, ORR, OS, complete remission, complete response, partial response, EFS, MRD and others • Safety endpoints reported in studies, including AEs, serious AEs, AEs leading to death, treatment discontinuations and others • HRQoL endpoints reported, including all PROs • HRQoL cancer specific: FACT-G, Mental Adjustment to Cancer Scale	Not including the outcome of interest
Study design	Randomised controlled trials • Non randomised trials, or single arm trials flagged only if the population and outcomes are of interest	• Not study type of interest • Not publication type of interest

AE, adverse event; EFS, event-free survival; FACT-G; functional assessment of cancer therapy-general; HRQoL, health-related quality of life; iNHL, indolent non-Hodgkin lymphoma; MRD, minimal residual disease; ORR, overall response rate; OS, overall survival; PFS, progression-free survival; PRO, patient-reported outcomes;

Furthermore, null entries (no information reported in the title or abstract), duplicates of existing entries, and abstracts and contents that have been reported elsewhere were excluded from the search.

Two additional exclusion criteria were included to filter the search results further:

1. At least one of the treatment arms includes treatment with MabThera (rituximab)
2. Studies assessing induction and maintenance treatment phases should not have a condition of successful completion of the induction treatment for patients in order to enter maintenance phase

4.1.3.2 Review strategy

Eligibility for inclusion was assessed using exclusion and inclusion criteria as detailed above. All citations were independently screened by two analysts, with any discrepancies resolved by discussion. A third reviewer was consulted for unresolved disagreements.

Once eligible publications were identified, full papers were obtained and screened again on the basis of the complete manuscript – rather than abstract only – to ensure eligibility. Identical eligibility criteria were used for both steps of the screening processes. As for the first step, two analysts conducted independent reviews of the full publications with a third reviewer consulted for any disagreements.

4.1.4 Search results. A flow diagram of the numbers of studies included and excluded at each stage should be provided using a validated statement for reporting systematic reviews and meta analyses, such as the PRISMA flow diagram. The total number of studies in the statement should equal the total number of studies listed in section 4.2.

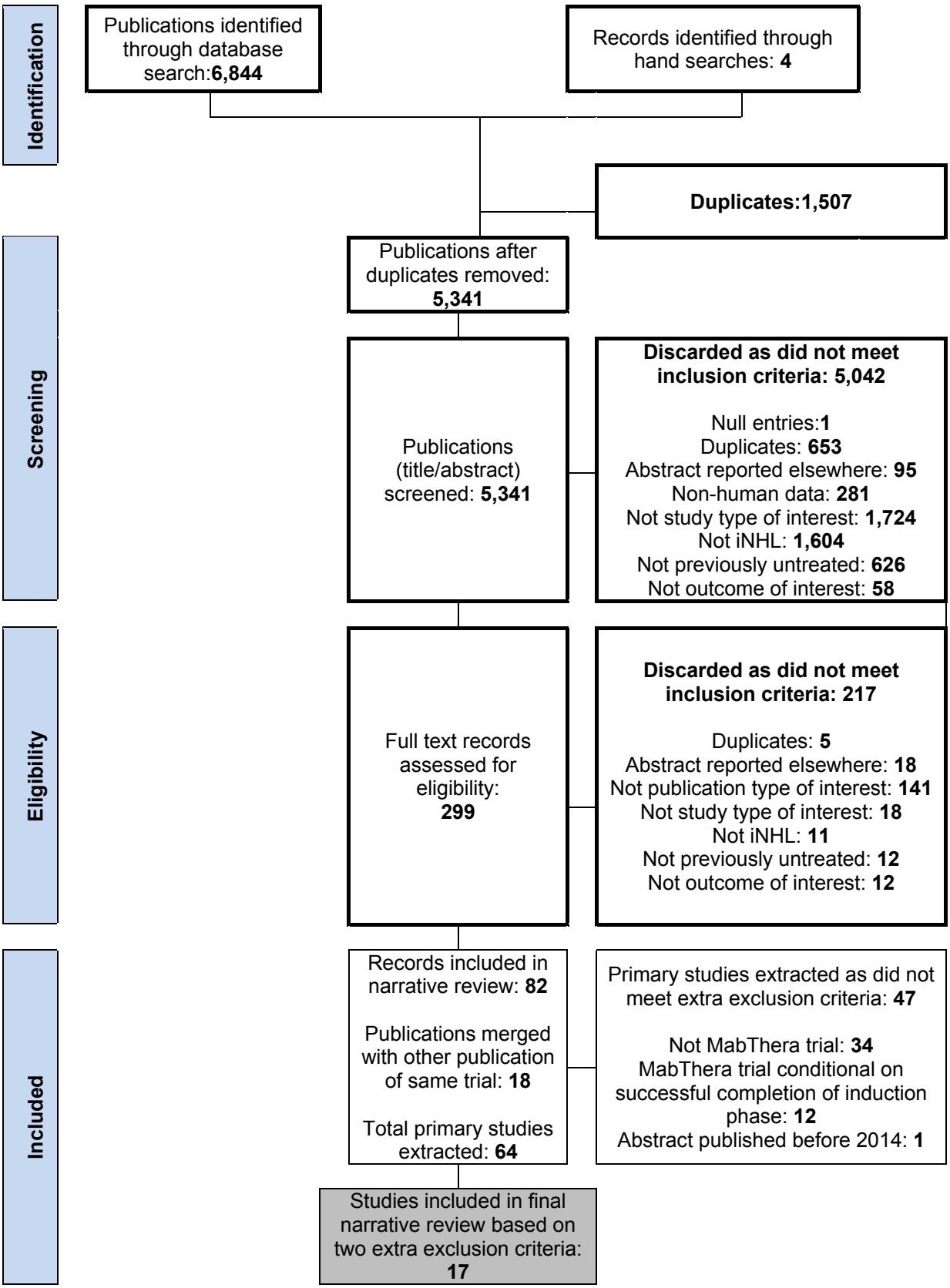
When data from a single study have been drawn from more than 1 source (for example, a poster and a published report) or when trials are linked (for example, an open label extension to a randomised controlled trial [RCT]), this should be clearly stated.

Provide a complete reference list for excluded studies in an appendix.

Database searches identified 6,844 studies with 4 additional publications identified through other sources. After removing duplicates and screening titles and abstracts, 299 articles were reviewed in full. In the end, 64 primary studies were included in the narrative synthesis, which was reduced further to 17 studies for the final narrative review following the addition of the two additional exclusion criteria

Of these 20 studies, one was found to be relevant to the decision problem in question. The 19 records excluded from the systematic review at the full-text review stage can be found in Appendix 4.

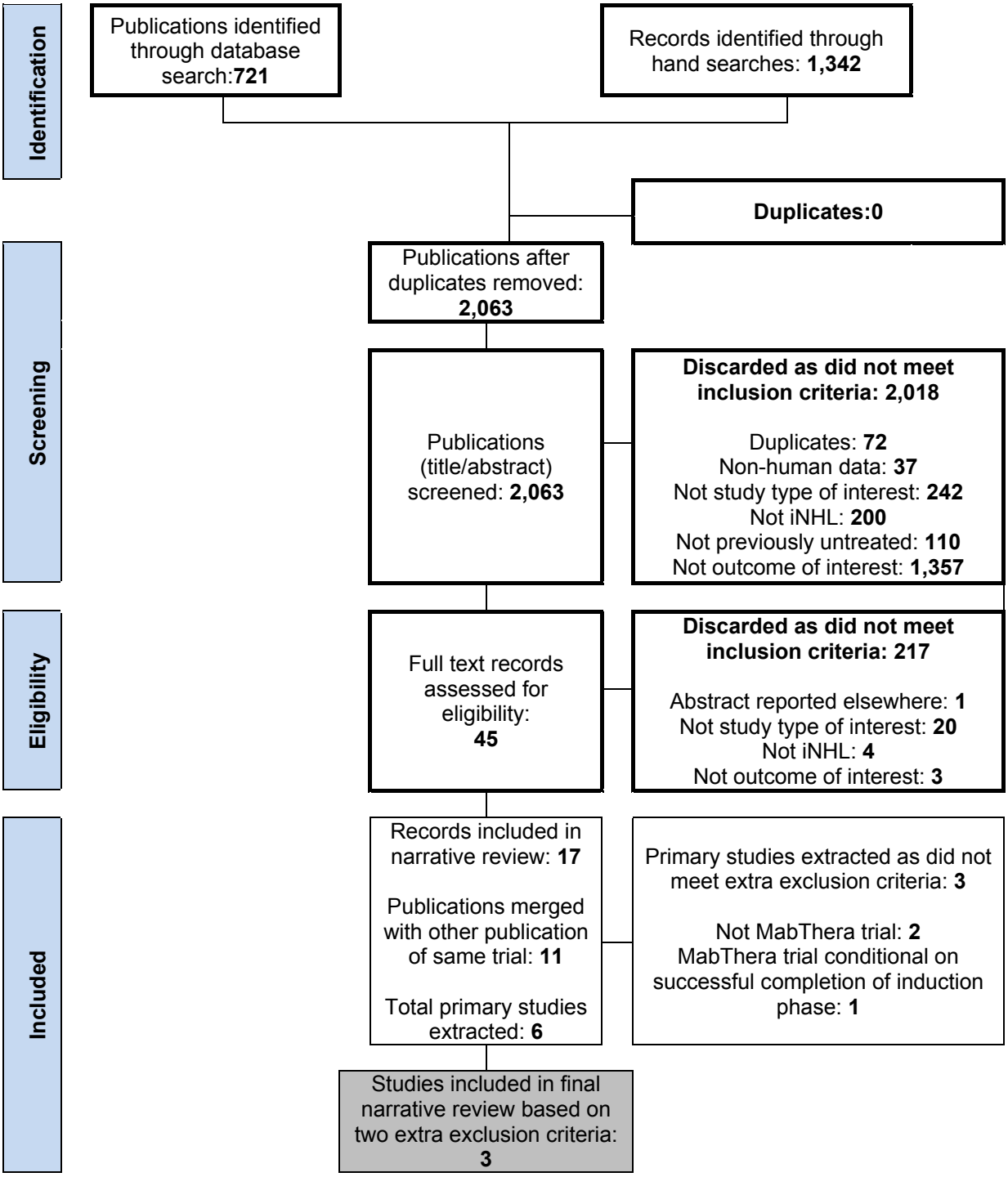
Figure 5: PRISMA flow diagram for clinical SLR (initial search)



The SLR was updated in March 2017, which identified an additional 2,063 citations. After removing duplicates and screening titles and abstracts, 45 articles were reviewed in full. In

the end, 3 primary studies were included in the narrative synthesis in addition to the 17 primary studies included previously.

Figure 6: PRISMA flow diagram for clinical SLR (updated search)



Of these 20 studies, one was found to be relevant to the decision problem in question. The 19 records excluded can be found in Appendix 4.

4.2 List of relevant randomised controlled trials

4.2.1 In a table, present the list of relevant RCTs comparing the intervention with other therapies (including placebo) in the relevant patient group. Highlight which studies compare the intervention directly with the appropriate comparator(s) with reference to the decision problem. If there are none, state this.

See Table 17.

4.2.2 When the RCTs listed above have been excluded from further discussion, justification should be provided to ensure that the rationale for doing so is transparent. For example, when RCTs have been identified, but there is no access to the level of data required, this should be stated.

Nomenclature used for the GALLIUM study in the clinical effectiveness section:

- Gazyvaro in combination with chemotherapy as induction therapy is abbreviated as G-chemo
- Gazyvaro monotherapy as maintenance is abbreviated as G
- The complete treatment regimen, i.e. induction plus maintenance therapy is therefore abbreviated as G-chemo+G. This represents the regimen as per the anticipated Marketing Authorisation

- MabThera in combination with chemotherapy as induction therapy is abbreviated as R-chemo
- MabThera monotherapy as maintenance is abbreviated as R
- The complete treatment regimen, i.e. induction plus maintenance therapy is therefore abbreviated as R-chemo+R.

Evidence for the clinical effectiveness, safety and tolerability of Gazyvaro in combination with chemotherapy as induction therapy, followed by Gazyvaro monotherapy maintenance has been demonstrated in one phase III, open-label randomised controlled trial, GALLIUM (BO21223, NCT01332968).(Clinical Trials.Gov)

This study compared Gazyvaro combined with chemotherapy (**G-chemo**) as induction, followed by Gazyvaro monotherapy as maintenance (**G**), with MabThera combined with chemotherapy (**R-chemo**) as induction followed by MabThera monotherapy as maintenance (**R**) in previously untreated patients with advanced iNHL.

Patients with iNHL were enrolled in the study, however, in accordance with the anticipated Marketing Authorisation, the data reported in this submission will focus on the subgroup of patients with FL. Subgroup analyses, stratified by demographics and baseline characteristics within this population will also be presented.

A summary of the GALLIUM clinical trial and available publications is provided in Table 17. No further randomised controlled trials comparing the efficacy or safety of Gazyvaro in combination with chemotherapy for the first-line treatment of FL were identified.

Table 17: List of relevant RCTs and publications

Trial number (name): NCT01332968 (GALLIUM)	
Sponsor	F. Hoffmann-La Roche, Ltd.
Intervention	G-chemo as induction followed by G maintenance monotherapy
Comparator	R-chemo as induction followed by R maintenance monotherapy
Population	• Previously untreated CD20-positive iNHL • FL (grade 1–3a) or splenic/nodal/extranodal MZL • Stage III/IV or stage II bulky disease (≥7cm) requiring treatment • Aged ≥18 years • ECOG 0– 2
Study references	**Primary analysis clinical cut-off date: 31st January, 2016** • Marcus R, et al. Oral presentation from Abstract 6, ASH 2016 (Marcus R et al., 2016) • Pott C, et al. Oral presentation from Abstract 613, ASH 2016 (MRD analysis)(Pott C et al., 2016) • BO21223 Primary Clinical Study Report (report number 1067980)(F. Hoffmann-La Roche Ltd, 2016) • Updated Gazyvaro SmPC (currently being revised) (F. Hoffmann-La Roche Ltd, 2017) • Primary manuscript to be published in New England Journal of Medicine

ASH, American Society of Hematology; ECOG, Eastern Cooperative Oncology Group; FL, follicular lymphoma; G-chemo, G with chemotherapy as induction; iNHL, indolent non-Hodgkin lymphoma; MRD, minimum residual disease; MZL, marginal zone lymphoma; R-chemo, MabThera with chemotherapy as induction; SmPC, Summary of Product Characteristics

This submission presents the results of the primary data analyses (clinical cut-off 31st January 2016), including all 1,401 patients randomised to the two study arms. Data from an updated analysis (clinical cut-off 16th September 2016) is presented where available.

4.3 Summary of methodology of the relevant randomised controlled trials

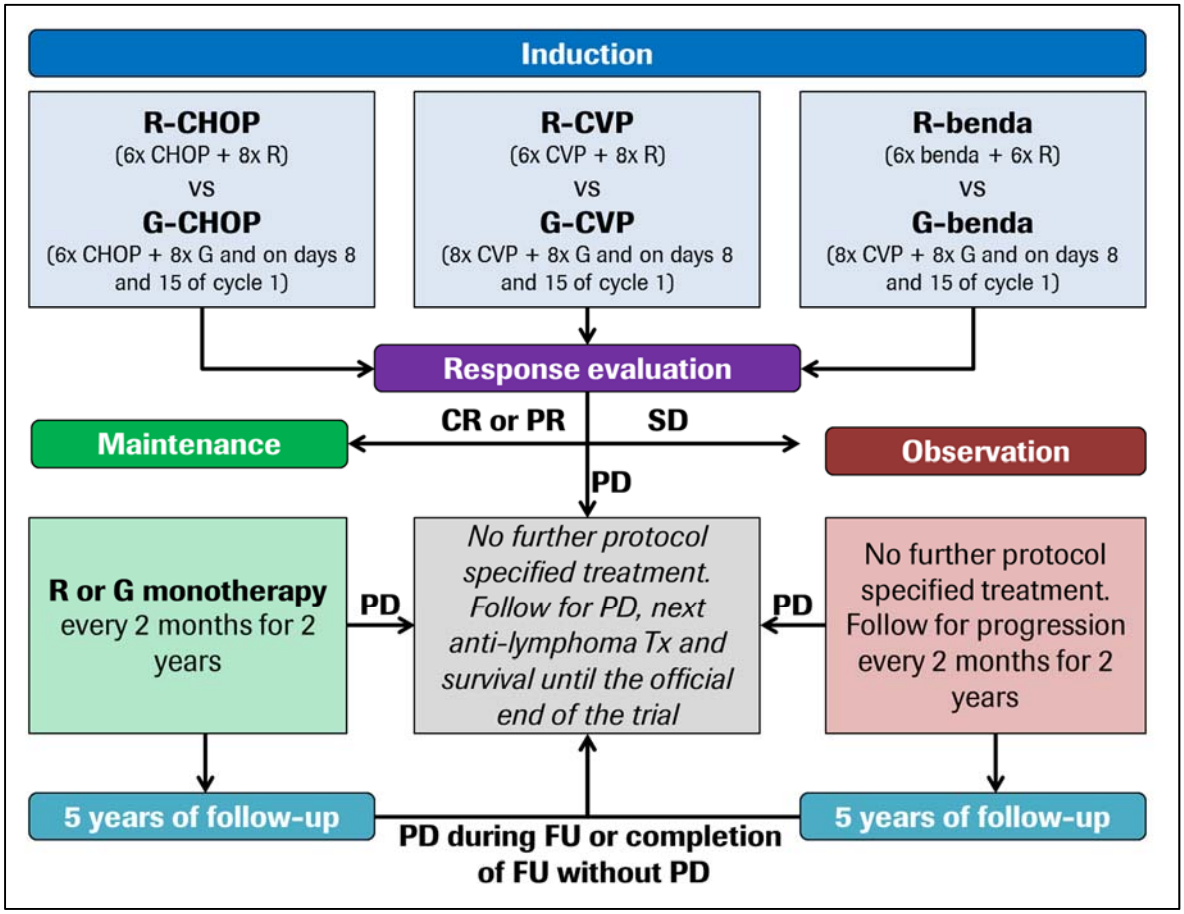
4.3.1 Items 3 to 6b of the CONSORT checklist should be provided for all RCTs listed:

Unless otherwise stated, all the information presented below is sourced from the primary clinical study report (data cut off 31st January 2016) (F. Hoffmann-La Roche Ltd, 2016).

Trial design

GALLIUM is an ongoing Phase III, open-label, multicentre, randomised study to investigate the efficacy and safety of G-chemo followed by G maintenance monotherapy for responders (complete response [CR] or partial response [PR]), compared with R-chemo followed by R-maintenance therapy for responders, in patients with previously untreated advanced indolent NHL requiring treatment.

Figure 7: GALLIUM study design schematic



CR, complete response; FU, follow-up; G, Gazyvaro; PD, progressive disease; PR, partial response; R, MabThera; SD, stable disease

Prior to the initiation of the study, each site chose one of three chemotherapy regimens (CHOP, CVP, or bendamustine) that was considered to be the standard of care for follicular

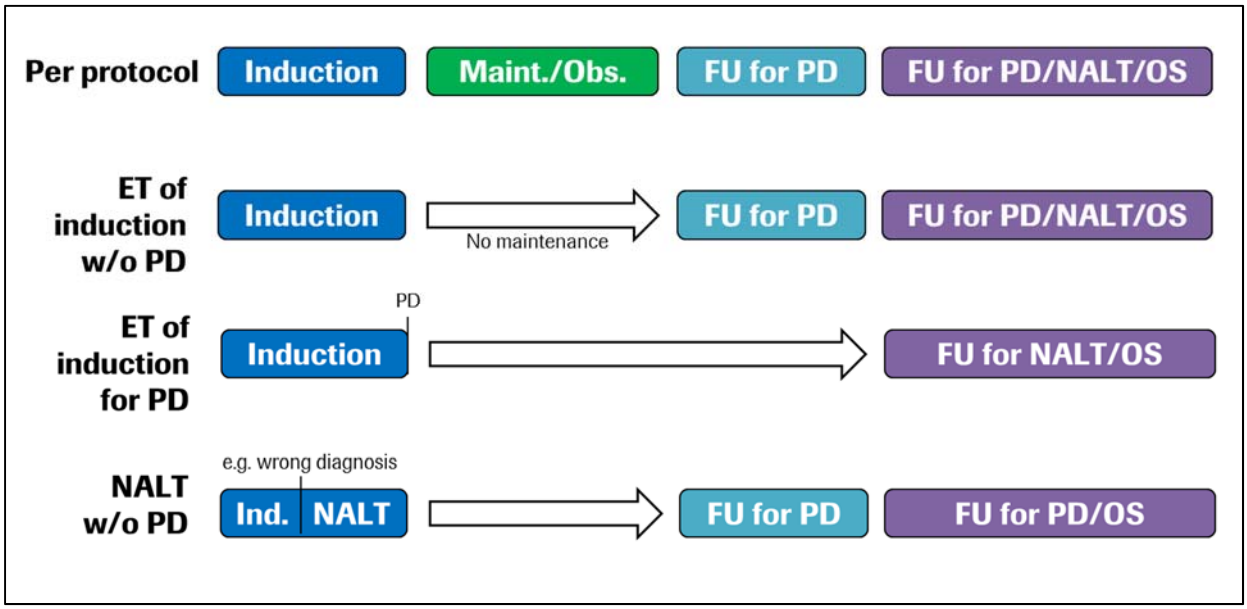
lymphoma; all patients with follicular lymphoma at that site received the chosen chemotherapy regimen for the duration of the study. For non-follicular NHL, the investigator had the option of choosing one of the three chemotherapy regimens (CHOP, CVP, or bendamustine) for each patient.

Approximately 1,200 patients with FL were recruited and randomly assigned in a 1:1 ratio to either G-chemo followed by G-maintenance in responders, or R-chemo followed by R-maintenance in responders. In addition, approximately 200 patients with MZL (splenic, nodal, or extranodal) were to be recruited and randomly assigned in a 1:1 ratio to the two treatment arms. Stratification factors for randomisation were:

- Chemotherapy regimen (CHOP, CVP or bendamustine)
- FLIPI (low or high for FL)
- Geographic region (Western Europe, Eastern Europe, South and Central America, North America, other

Following the completion of induction therapy, patients received maintenance therapy (if they achieved a CR or PR) or underwent observation (patients with stable disease [SD]), and were followed clinically every 2 months for 2 years. For patients who had not progressed at the maintenance or observation completion visit, disease assessments continued every 3 months for 3 years then every 6 months for 2 years until disease progression. After 5 years of follow-up or disease progression (whichever came first), patients were then followed every 6 months for OS and new anti-lymphoma treatment (NALT), or for disease progression if applicable, until the end of the study, which is estimated as 10.2 years after inclusion of the first patient. Patients who terminated early without PD were followed for PD, and in the extended follow-up for PD, NALT and OS. Patients who terminated induction early because of PD went directly into the extended follow-up for NALT and OS. Patients who discontinued the protocol-defined treatment path and needed to start a NALT in the absence of disease progression (e.g., if wrong diagnosis at screening and new diagnosis required a change of treatment) were followed for disease progression and OS. The study phases in GALLIUM are summarised below.

Figure 8: GALLIUM study phases



ET, early-termination; FU, follow-up; Ind., induction; Maint., maintenance; NALT, new anti-lymphoma treatment; Obs., observation; OS, overall survival; PD, progressive disease

An Independent Data Monitoring Committee (IDMC) was established to monitor patient safety as well as perform the interim analyses: two for futility and one for efficacy; see Section 4.4 for further details.

All patients were assessed for disease response by the investigator through the use of regular clinical and laboratory examinations and computed tomography (CT) or magnetic resonance imaging (MRI) scans according to a modified version of the Revised Response Criteria for Malignant Lymphoma (Cheson BD et al., 2007). An independent radiologic and oncologic review of the responses of all patients by an IRC was also conducted for the futility and efficacy interim analyses (including overall response rate [ORR] with and without PET).

Evaluation of response by FDG-PET was mandatory at induction completion/end-of-therapy visit (only if screening PET was positive) 6–8 weeks after Day 1 of the last cycle of induction for the first 170 patients with follicular lymphoma at those sites that had a PET scanner available. Additional optional FDG-PET scans (e.g., in all other lymphoma patients beyond the first 170 patients with FL) were permitted if the investigator chose to perform them. In the overall study population, FDG-PET was optional upon investigator's discretion. Assessment of PET results was to be made according to the criteria established by the Imaging Subcommittee of International Harmonization Project in Lymphoma (Juweid et al., 2007).

Eligibility criteria

GALLIUM included adult patients with previously untreated advanced indolent NHL; the specific inclusion and exclusion criteria are detailed in Table 18 and Table 19 below, respectively.

Table 18: GALLIUM inclusion criteria

Inclusion criteria
• Age ≥ 18 years • ECOG performance status of 0, 1 or 2 • Histologically documented, CD20-positive, indolent B-cell consisting of one of the following: FL (Grades 1–3a), splenic MZL, nodal MZL, or extranodal MZL • Stage III or IV disease or Stage II bulky disease (bulky disease defined as tumour diameter ≥7 cm) • For patients with FL, requirement for treatment defined as meeting one of the following criteria: o Bulky disease (nodal or extranodal mass ≥7 cm in the greatest diameter o Local symptoms or compromise of normal organ function due to progressive nodal disease or extranodal tumour mass o Presence of B symptoms (fever, drenching night sweats, or unintentional weight loss of >10% body weight over a period of 6 months or less) o Presence of symptomatic extranodal disease (e.g., pleural effusions, peritoneal ascites) o Cytopenias due to underlying lymphoma (i.e., absolute neutrophil count <1.0 x 10^9/L, haemoglobin <10 g/dL, and/or platelet count < 100 x 10^9/L) o Involvement of ≥3 nodal sites, each with a diameter of ≥3 cm o Symptomatic splenic enlargement • At least one bi-dimensionally measurable lesion (>2 cm in its largest dimension by CT scan or magnetic resonance imaging [MRI]) • Adequate haematologic function (unless abnormalities are related to NHL), defined as follows: o Haemoglobin ≥9.0 g/dL o Absolute neutrophil count ≥1.5x10^9/L o Platelet count ≥75x10^9/L • Able and willing to provide written informed consent and to comply with the study protocol

CT computed tomography; ECOG, Eastern Cooperative Oncology Group; FL, follicular lymphoma; MRI, magnetic resonance imaging; MZL, marginal zone lymphoma;

Table 19: Key GALLIUM exclusion criteria

Exclusion criteria
• History of severe allergic or anaphylactic reactions to monoclonal antibody therapy (e.g., patients in whom dosing with rituximab would be contraindicated for safety reasons) • Central nervous system lymphoma, leptomeningeal lymphoma, or histologic evidence of transformation to a high-grade or diffuse large B-cell lymphoma • Grade 3b FL, SLL, or WM

- For patients with FL: prior treatment for NHL by chemotherapy, immunotherapy, or radiotherapy
    - Low-dose methotrexate (MTX) in rheumatoid arthritis (typically 7.5 mg to a maximum of 20 mg/week) was not considered chemotherapy for lymphoma. It was recommended to stop MTX 2-3 weeks prior to starting immunochemotherapy since the combination of MTX and immunochemotherapy increases the risk of immunosuppression and of infection
- Regular treatment with corticosteroids during the 4 weeks prior to the start of Cycle 1, unless administered for indications other than NHL at a dose equivalent to ≤30 mg/day prednisone
- Patients with a history of confirmed progressive multifocal leukoencephalopathy (PML)
- History of prior other malignancy with the exception of:
    - Curatively treated basal cell carcinoma or squamous cell carcinoma of the skin or carcinoma in situ of the cervix at any time prior to study
    - Other cancers not specified above which have been curatively treated by surgery alone and from which subject is disease-free for ≥5 years without further treatment
- Evidence of significant, uncontrolled concomitant diseases that could affect compliance with the protocol or interpretation of results, including significant cardiovascular disease (such as New York Heart Association Class III or IV cardiac disease, severe arrhythmia, myocardial infarction within the previous 6 months, unstable arrhythmias, or unstable angina) or pulmonary disease (including obstructive pulmonary disease and history of bronchospasm)
- For patients who received CHOP: LVEF <50% by MUGA scan or echocardiogram
- Any of the following abnormal laboratory values:
    - Creatinine >1.5 x ULN or creatinine clearance <40 mL/min
    - AST or ALT >2.5 x ULN
    - Total bilirubin >1.5 x ULN (or >3 x ULN for patients with document Gilbert syndrome)
    - International normalised ratio (INR) or prothrombin time (PT) > 1.5 x ULN in the absence of therapeutic anticoagulation
    - Partial thromboplastin time (PTT) or activated PTT (aPTT) >1.5 x ULN in the absence of a lupus anticoagulant

ALT, alanine transaminase; AST, aspartate transaminase; CHOP, cyclophosphamide, doxorubicin, vincristine, prednisone; INR, international normalised ratio; LVEF, left ventricular ejection fraction; MTX, methotrexate; MUGA, multigated radionuclide angiography; PML, progressive multifocal leukoencephalopathy; (a)PTT, (activated) partial thromboplastin time; PT, prothrombin time; SLL, small lymphocytic lymphoma; ULN, upper limit of normal; WM, Waldenstrom macroglobulinaemia

Settings and locations of data collection

A total of 1401 patients were randomised in the study (699 patients to the R-chemo arm and 702 patients to the G-chemo arm). The first patient was randomised on 6th July, 2011 and the last patient on 5th February, 2014. Patients were recruited from 177 investigational sites in 18 countries and the highest recruiting countries were United Kingdom (293 patients), Germany (237 patients), Canada (138 patients), Australia (136 patients), and Japan (129 patients).

Trial drugs and concomitant medications

Patients were randomised in a 1:1 fashion through an interactive voice response system (IVRS) to the G-chemo or R-chemo arms. Randomisation occurred separately for the patients with FL and MZL.

R-chemo was chosen as the control arm for the study as it was recommended in expert international treatment guidelines (the 2009 ESMO Guidelines Working Group recommendations (Dreyling, 2009) and the 2010 NCCN guidelines (Zelenetz et al., 2010)) for newly diagnosed patients with bulky stage II, stage III or stage IV disease requiring treatment at the time of initiation of the study. Furthermore, two randomised clinical trials demonstrated the benefit of R-maintenance versus observation in responding patients with both previously untreated and relapsed FL (Salles et al., 2011, van Oers et al., 2006); these studies set a new standard consisting of immunochemotherapy induction followed by R-maintenance for the treatment of patients with previously untreated advanced FL.

G-chemo followed by G monotherapy maintenance

In the G-chemo arm, eight to ten doses of Gazyvaro at 1000 mg were administered by IV infusion with the accompanying chemotherapy regimen during induction.

- **G-CHOP:** G was administered on Days 1, 8, and 15 of Cycle 1 and on Day 1 of Cycles 2–8 (21-day cycles). CHOP was administered on Day 1, with prednisone/prednisolone/methylprednisolone also administered on Days 2–5 of Cycles 1–6
- **G-CVP:** G was administered on Days 1, 8, and 15 of Cycle 1 and on Day 1 of Cycles 2–8 (21-day cycles). CVP was administered on Day 1, with prednisone/prednisolone/methylprednisolone also administered on Days 2–5 of Cycles 1–8
- **G-bendamustine:** G was administered on Days 1, 8, and 15 of Cycle 1 and on Day 1 of Cycles 2–6 (28-day cycles). Bendamustine was administered on Days 1 and 2 of Cycles 1–6, with prednisone/prednisolone/methylprednisolone administered on Day 1 of Cycle 1

Patients randomised to receive G-chemo who achieved a CR or PR at the end of induction therapy continued to receive G-maintenance at 1000 mg every 2 months until disease progression, or for 2 years.

R-chemo followed by R monotherapy maintenance (control arm)

In the R-chemo arm, six to eight doses of R at 375 mg/m^2 were administered by IV infusion with the accompanying chemotherapy regimen during induction.

- **R-CHOP:** R was administered on Day 1 of Cycles 1–8 (21-day cycles). CHOP was administered on Day 1, with prednisone/prednisolone/methylprednisolone also administered on Days 2–5, of Cycles 1–6
- **R-CVP:** R was administered on Day 1 of Cycles 1–8 (21-day cycles). CVP was administered on Day 1, with prednisone/prednisolone/methylprednisolone also administered on Days 2–5, of Cycles 1–8
- **R-bendamustine:** R was administered on Day 1 of Cycles 1–6 (28-day cycles). Bendamustine was administered on Days 1 and 2 of Cycles 1–6, with prednisone/prednisolone/methylprednisolone also administered on Day 1 of Cycle 1.

Patients randomised to receive R-chemo who achieved a CR or PR at the end of induction therapy continued to receive R-maintenance at 375 mg/m^2 every 2 months until disease progression, or for 2 years.

CHOP

It was planned to administer a total of six 21-day cycles of CHOP. If CHOP was discontinued for any reason other than toxicity, patients were to be discontinued from study treatment and entered follow-up directly without maintenance. CHOP was administered according to the standard preparation and infusion procedures of each investigational site.

CVP

It was planned to administer a total of eight 21-day cycles of CVP. If CVP was discontinued for any reason other than toxicity, patients were to be discontinued from study treatment and entered follow-up directly without maintenance. CVP was administered according to the standard preparation and infusion procedures of each investigational site.

Bendamustine

It was planned to administer a total of six 28-day cycles of bendamustine. If bendamustine was discontinued for any reason other than toxicity, patients were to be discontinued from study treatment and entered follow-up directly without maintenance.

Premedication

All G or R infusions were administered after premedication with oral acetaminophen/ paracetamol and an antihistamine. Patients who were considered to have a high tumour burden and who were considered at risk for tumour lysis by the investigator also received tumour lysis prophylaxis prior to the initiation of treatment.

Cyclophosphamide, doxorubicin, and bendamustine have a moderate risk of emesis, therefore it was recommended that infusions be administered following premedication with a serotonin antagonist or per institutional practice.

Concomitant therapies

The permitted and prohibited concomitant medications for GALLIUM are detailed below.

Table 20: Permitted concomitant medications

Permitted medications
• Any prescription medicine or over-the-counter preparations used by a patient between the 7 days preceding the study entry evaluation and the end of study visits • Oral contraceptives, hormone-replacement therapy, or other maintenance therapy • Prophylactic anti-viral medication to prevent hepatitis B reactivation • Rasburicase for the treatment of tumour lysis syndrome and the prevention of hyperuricaemia was allowed according to institutional guidelines • Antibiotic and/or anti-viral prophylaxis according to institutional guidelines • Primary prophylaxis with granulocyte colony stimulating factors (G-CSFs) was recommended as per the ASCO, EORTC, and ESMO guidelines, namely, in patients who were ≥60 years of age and/or with comorbidities (the use of G-CSF prophylaxis was strongly recommended in Cycle 1 for all patients treated with G-CHOP) • Harvesting of stem cells by G-CSF alone (no additional chemotherapeutic agent) was allowed only if it was done between Cycle 5 Day 1 and Cycle 8 Day 1 (R/G-CHOP or R/G-CVP) or Cycle 4 Day 1 and Cycle 6 Day 1 (R/G-Bendamustine) • Acetaminophen (≥ 500 mg) and/or H1− and H2-histamine-receptor antagonists (e.g., diphenhydramine, ranitidine) for the symptomatic treatment of Gazyvaro infusion-related temperature elevations of > 38.5°C or other minor infusion-related symptoms • Additional supportive therapies (e.g., supplemental oxygen, β2-agonists/epinephrine, and/or corticosteroids) for the treatment of serious infusion-related events manifested by dyspnoea, hypotension, wheezing, bronchospasm, tachycardia, reduced oxygen saturation, or respiratory distress • Myeloid growth factors for the primary prevention and treatment of febrile neutropenia • Mesna as prophylaxis for haemorrhagic cystitis per institutional guidelines for patients treated

with CHOP or CVP

Table 21: Prohibited concomitant medications

Prohibited medications
• Cytotoxic chemotherapy (other than bendamustine, cyclophosphamide, doxorubicin, or vincristine) • Radiotherapy • Immunotherapy (other than G and R) • Hormone therapy (other than contraceptives, hormone-replacement therapy, or megestrol acetate) • Any therapies intended for the treatment of NHL whether FDA approved or experimental (outside of the study)

Study endpoints

Study endpoints and study period when data were collected are summarised in Table 22.

Primary endpoint

- **Progression-free survival (PFS) in patients with FL**, was defined as the time from randomisation to the first occurrence of progression or relapse as assessed by the investigator according to the Revised Response Criteria for Malignant Lymphoma (Cheson BD et al., 2007) or death from any cause

Secondary endpoints

The following secondary outcome measure applied to patients with previously untreated advanced indolent NHL (i.e., overall population, ITT):

- **Investigator-assessed PFS**

The following secondary outcome measures applied to patients with previously untreated advanced indolent NHL (i.e., ITT population) and to the subset of patients with previously untreated advanced FL:

- **Independent-review committee (IRC) assessed PFS**
- **CR and overall response** (CR or PR) at the end of induction (in the FL and ITT population), as assessed by the:
    - Investigator with and without FDG-PET
    - IRC with and without FDG-PET

- **Overall survival (OS)** (in the FL and ITT population) defined as the time from randomisation to death from any cause
- **Event-free survival (EFS)**, defined as the time from randomisation to disease progression/relapse as assessed by the investigator, death from any cause, or start of the next anti-lymphoma treatment (NALT)
- **Disease-free survival (DFS)**, defined for patients with a best overall response (BOR) of CR as the time from first occurrence of a documented CR to PD as assessed by the investigator or death from any cause. Patients who have had no documented disease progression or have not died after CR were censored at the last disease assessment date
- **Duration of response (DoR),** defined for patients with a BOR of CR or PR as the time from first occurrence of a documented CR or PR to disease progression/relapse as assessed by the investigator or death from any cause. For patients achieving a response who have not progressed, relapsed, or died at the time of the analysis, duration of response will be censored on the date of last disease assessment.
- **Time to NALT**, defined as the time from randomisation to start of new non-protocol anti-lymphoma therapy or death from any cause

Health-related patient-reported outcomes (PROs)

- Change from baseline to the end of study in patient-reported outcomes (PROs) based on the FACT-Lym instrument, as outlined below:
    - Change from baseline in all domains of the FACT-G
    - Change from baseline in the total outcome index (TOI) (range, 0−116): sum of physical well-being (7 items), functional well-being (7 items), and Lym subscale (15 items) scores
    - Change from baseline in the FACT-Lym subscale score (range, 0−60): 15 lymphoma-specific items
    - Change from baseline in the FACT-Lym total score (range, 0−168): sum of physical well-being (7 items), social/family well-being (7 items), emotional well-being (6 items), functional well-being (7 items), and Lym subscale (15 items) scores
- EQ-5D summary scores at baseline, during treatment, after treatment, at the last assessment prior to progression, and at the first assessment after progression

Exploratory efficacy endpoints

- To assess the prognostic and predictive value of BCL2/IgH rearrangement and other markers of minimum residual disease (MRD) in patients with FL at baseline, during induction, at the completion of induction therapy, during maintenance therapy or observation, and during follow-up (IgH clonality could be used as a marker of MRD in patients without an identifiable BCL2/IgH translocation at baseline)
- End-of-maintenance response in FL patients (including all patients randomised other than those who had not yet reached the end of the maintenance assessment)

Table 22: Summary of endpoints and study period when data were collected

Induction period only	**All study periods (induction + maintenance + follow up)**
Complete response	Progression-free survival (investigator and IRC)
End-of-treatment overall response	Overall survival
	Best overall response
	Disease-free survival
	Event-free survival
	Duration of response
	Patient-reported outcomes
	Minimal residual disease
	End-of-maintenance response

Subgroup analyses

Subgroup analyses of investigator-assessed PFS, IRC-assessed PFS, CR rate, and ORR (all without PET) were planned for the FL and ITT populations for each of the following:

- Stratification factors (chemotherapy regimen, FLIPI or IPI risk group, geographic region)
- Age at randomisation
- Baseline characteristics and disease demographics (including but not limited to gender, race, ECOG performance status, Ann Arbor stage)

Safety reporting and analyses

Safety analyses were conducted for the whole study period and selected safety analyses by study phase. All safety analyses were performed on the safety population within all

randomised patients and the FL subsets. Safety assessments included adverse events (AEs) (including serious adverse events [SAEs]), standard laboratory assessments, and vital signs.

After the initiation of the study medication, AEs and SAEs were recorded as follows (until patient began NALT):

- All AEs (related and unrelated) were recorded up to 28 days are the last dose of study drug
- Grade ≥3 AEs (related and unrelated) were recorded up to 6 months after the last dose of study drug
- Grade 3 or 4 infections (related and unrelated) were recorded up to 24 months after the last dose of study drug
- Unrelated SAEs were recorded up to 12 months after the last dose of study drug
- Study drug-related SAEs were recorded indefinitely (even if the study had been closed).

Adverse events of particular interest (AEPIs) included all events of special interest and additionally, all events for which a separate analysis has been performed. AEPIs were defined prior to the primary analysis based on the mode of action of G and the need to gather further safety information. Serious events of infusion-related reactions (IRRs), neutropenia and infection, and all cases of TLS were considered AEs of special interest (AESI), as well as being AEPIs. AESIs had to be reported by the investigator to the Sponsor within 24 hours of learning of the event.

Table 23: Adverse events of particular interest and special interest

AEPIs
• IRR - AEs related to any study medication (not specific to Gazyvaro), which occurred during infusion or within 24 hours from the end of infusion • Neutropenic events • Prolonged neutropenia - initial absolute neutrophil count (ANC) <1.0 x 10^9/L following last antibody administration (LAA) and ANC <1.0 x 10^9/L at last previous visit before LAA • Late onset neutropenia - initial ANC <1.0 x 10^9/L following LAA and ANC within normal range (≥1.0 x 10^9/L) at last previous visit before LAA • All infections • All TLS events • Thrombocytopenia • Acute thrombocytopenia – occurring during or within 24 hours post infusion

• GI perforation • Cardiac events • Secondary malignancy – defined by any neoplasms benign, malignant and unspecified (including cysts and polyps) starting 6 months after the first study drug intake • Hepatitis B reactivation - an elevation of HBV DNA post baseline (HBV DNA ≥ 100 IU/mL) maintained for two consecutive assessments, using central laboratory results
AESIs
• Serious IRRs - SAEs related to any study medication (not specific to G), which occurred during infusion or within 24 hours from the end of infusion • Neutropenic events (serious) • Infections (serious) • TLS (all grades and irrespective of seriousness)

AEs, adverse events; AEPIs, adverse events of particular interest; AESI, adverse events of special interest; ANC, absolute neutrophil count; GI, gastrointestinal; HBV, hepatitis B virus; IRR, infusion-related reactions; LAA, last antibody administration; SAEs, serious adverse events; TLS, tumour lysis syndrome

4.4 Statistical analysis and definition of study groups in the relevant randomised controlled trials

4.4.1 During completion of this section consider items 7a (sample size), 7b (interim analyses and stopping guidelines), 12a (statistical methods used to compare groups for primary and secondary outcomes) and 12b (methods for additional analyses, such as subgroup analyses and adjusted analyses) of the CONSORT checklist.

Unless otherwise stated, all the information presented below is sourced from the GALLIUM primary CSR, data cut off 31st January 2016 (F. Hoffmann-La Roche Ltd, 2016). Analyses are relevant to both the ITT population and FL subset.

Analysis timing

Three interim analyses were planned: two for futility (one on CR and one on PFS) and one for efficacy (on PFS). The first interim analysis was based on differences in end-of-induction CR rates in the first 170 enrolled patients with FL. The IDMC reviewed the data on 24th October 2012 and recommended that the study continue.

The second interim analysis (futility on PFS) was conducted when 30% of the required investigator-assessed PFS events (i.e., approximately 111 events) had occurred. The clinical data cut-off for the second interim analysis was 20th February 2014. The IDMC reviewed the data on 31st July 2014 and recommended that the study continue.

The third interim analysis (efficacy) was planned after 67% of the events had occurred (i.e. approximately 248 events), and all patients had been enrolled and followed for an estimated minimum of 11 months. The clinical cut-off date for the third interim analysis was 31st January 2016. This analysis is now referred to as the primary analysis. The IDMC reviewed the data on 20th May 2016 and recommended that the study be fully analysed at this time, as the primary endpoint had been met.

4.4.2 For each trial listed, provide details of the trial population included in the primary analysis of the primary outcome and methods used to take account of missing data (for example, a description of the intention to treat analysis carried out, including censoring methods, or whether a per protocol analysis was carried out).

GALLIUM analysis populations

ITT FL population

The primary efficacy analysis population is the ITT FL population, defined as all randomised patients with follicular histology. Efficacy analyses were conducted according to the ITT principle, where patients were grouped according to their randomised treatment arm regardless of what treatments were actually received.

ITT – overall population

The primary and key secondary efficacy parameters were also determined in the overall ITT population, defined as all randomised patients.

Safety Population

The safety analysis population included all patients who received any amount of study drug (G, R, or chemotherapy [CHOP, CVP, or bendamustine]), and patients were analysed according to the treatment received (i.e., a patient who received G at least once for any reason was analysed under the G-chemo treatment arm; if only chemotherapy and/or R was received, the patient was analysed under the R-chemo treatment arm).

PET evaluable population

The “PET evaluable” subset contains all patients for whom the answer to the question “Were there any PET-avid lesions representing lymphoma?” on PET scan eCRF at baseline was “Yes”.

Patient reported outcomes

The PRO analyses included all randomised patients who had a baseline and at least one post-baseline PRO assessment. Patients in this subset were analysed according to their randomised treatment assignment, irrespective of the treatment received. The analyses were performed separately for FL and overall populations. PRO assessment was scheduled to continue up to and including the first assessment after disease progression was reported.

Subgroup analyses

Subgroup analyses of investigator-assessed PFS, IRC-assessed PFS, CR rate, and ORR (all without PET) were planned for the FL and overall populations according to prognostic factors to assess internal consistency.

The estimated probabilities in yearly intervals, as well as the hazard ratio and their 95% confidence intervals (for time-to-event endpoints) or response rates, as well as the odds ratio and their 95% confidence intervals (for binary endpoints), are reported separately for each level subgroup.

Sensitivity analysis

The FL population was the primary population for all efficacy sensitivity analyses.

The following sensitivity analyses for both IRC and investigator-assessed PFS were performed:

- Unstratified log-rank test
- Re-randomization test of the primary endpoint to assess the sensitivity of the stratified log-rank test to the dynamic randomisation procedure
- The impact of loss to follow-up was assessed by a worst-case analysis that assigns event outcomes to patients who withdrew prior to disease progression in the G arm at the next scheduled disease assessment date and censored outcomes to patients in the R arm at the last disease assessment date
- A missed assessment potential impact analysis was performed to assess the robustness of the result of the analysis of PFS. In this analysis, if patients missed an assessment prior to the date of the clinical data cut-off or prior to PD, they were

counted as having progressed of the day after their last complete response assessment

- PFS analyses were repeated with censoring at the initiation of NALT prior to disease progression, to assess potential confounding of the treatment effect estimates by subsequent therapy
- Patients who discontinued the study treatment for other reasons than disease progression or death were counted as having progressed at the time of discontinuation (event was date of last dose for early treatment discontinuations)
- Patients who died more than 6 months after their last response assessment and showed no sign of progression were censored at the last available response assessment.

4.4.3 For each trial, provide details of the statistical tests used in the primary analysis. Also provide details of the primary hypothesis or hypotheses under consideration, the power of the trial and a description of sample size calculation, including rationale and assumptions in a table. If the outcomes were adjusted for covariates, provide the rationale.

PFS was the primary efficacy endpoint of GALLIUM, defined as the time from the day of randomisation until the first documented day of disease progression, symptomatic deterioration, disease transformation, or death from any cause, whichever occurred first. Patients who did not experience documented disease progression or death were censored at the last valid (SD, PR, CR) tumour assessment prior to the clinical cut-off date.

PFS was compared using a two-sided log-rank test stratified by chemotherapy regimen (CHOP, CVP, or bendamustine), FL international prognostic index (FLIPI) risk group (low, intermediate, or high) in patients with FL or international prognostic index (IPI) risk group (low or low-intermediate vs. high-intermediate or high) in patients with non-follicular lymphoma.

The primary analysis of the study tested the equality of PFS distributions in the R-chemo and G-chemo arms the following null hypothesis with use of a two-sided stratified log rank test at an overall 5% significance level:

- Equality of PFS distributions in the G-chemo and R-chemo arms in the FL population by investigator assessment:

$$H_0: PFS_{G\text{-}chemo} = PFS_{R\text{-}chemo} \text{ versus } H_1: S_{G\text{-}chemo} \neq S_{R\text{-}chemo}$$

In the FL subset, estimates on the number of events required to demonstrate efficacy with respect to PFS were based on the following assumptions:

- Two-sided log rank test at the 0.05 level of significance
- Powered for the FL population
- Eighty percent power to detect a hazard ratio (HR) for G-chemo versus R-chemo of 0.74, corresponding to an improvement in 3-year PFS from 70.7% to 77.4% or in median PFS from 6 years to 8.1 years (35%)[4]
- Exponential distribution of PFS
- An annual dropout rate of 2.5%.

With the above assumptions, 370 PFS events were required to achieve 80% power for the primary analysis. Recruitment was staggered in order to recruit the first 170 patients at a smaller number of sites, followed by the activation of all sites after the IDMC meeting for futility based on CR rates. It was expected that during the first stage, after a 6-month ramp up, 18 patients per month would be recruited, and after the IDMC meeting and another 4-month ramp up, an accrual rate of 37 patients per month was expected.

The 1200 patients with FL enrolled over 49 months and followed for an additional 29 months after randomisation of the last patients were required to provide 370 PFS events, with a total duration for PFS follow-up estimated at approximately 78 months (6.5 years).

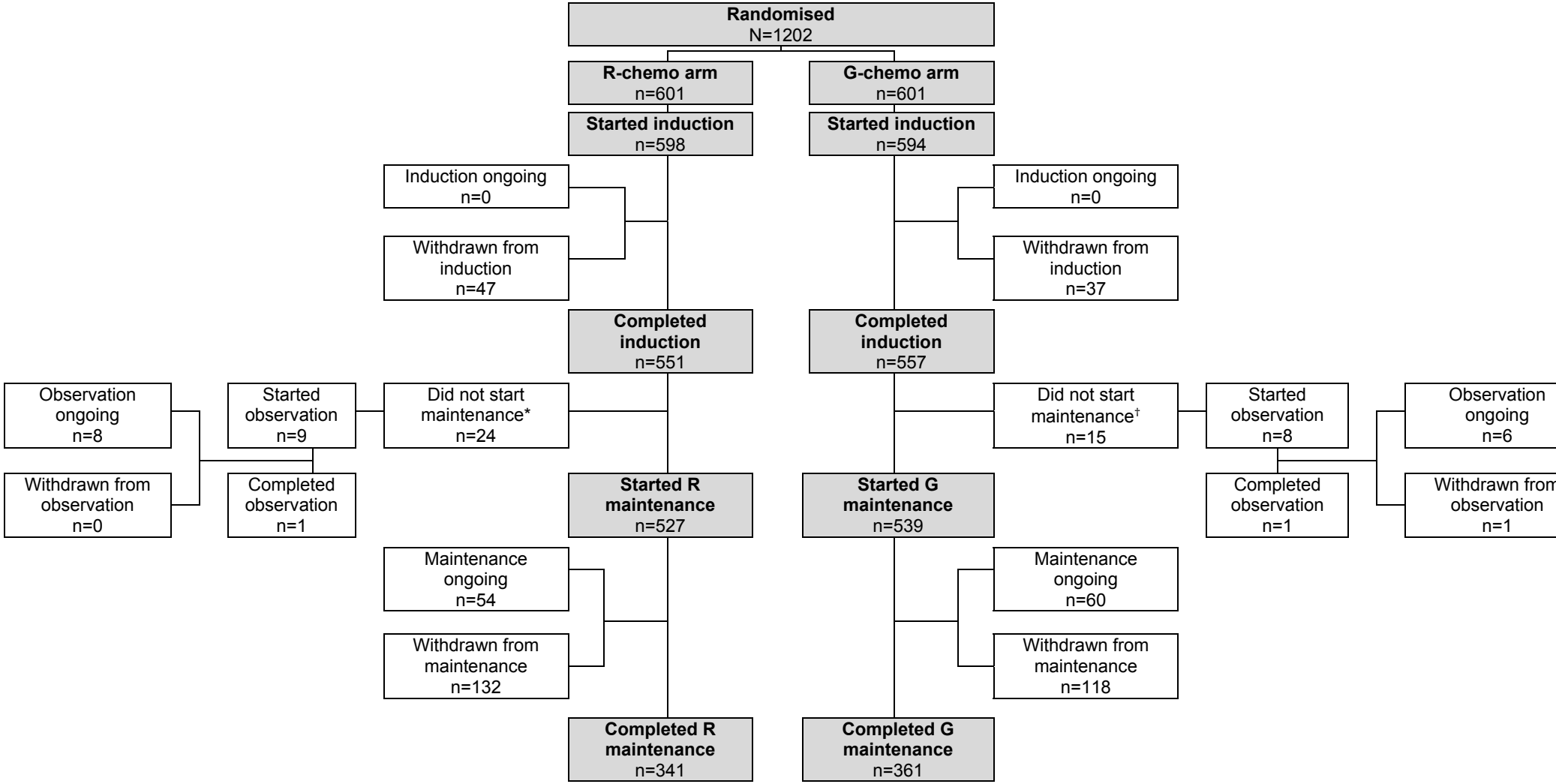
4.5 Participant flow in the relevant randomised controlled trials

4.5.1 Provide details of the numbers of participants who were eligible to enter the trials. Include the number of participants randomised and allocated to each treatment. Provide details of and the rationale for participants who crossed over treatment groups, were lost to follow up or withdrew from the RCT. Provide a CONSORT diagram showing the flow of participants through each stage of each of the trials.

A total of 1202 FL patients were randomised in the study (601 patients to the R-chemo arm and 601 patients to the G-chemo arm). Patient disposition for patients with FL at the clinical cut-off date is summarised below.

[4]estimates of median PFS were not likely to be reached in either study arm at either interim or final PFS analysis

Figure 9: Patient disposition (FL-ITT) clinical cut-off 31st January 2016



Reference: (F. Hoffmann-La Roche Ltd, 2016)

*24 patients did not start R-maintenance treatment due to: progressive disease between induction and maintenance (n=10); started observation (i.e., stable disease) (n=9); withdrawal by subject (n=3); physician decision (n=1); and other (n=1). †19 patients did not start G-maintenance treatment due to: progressive disease between induction and maintenance (n =10; started observation (i.e., stable disease) (n=8); and withdrawal by subject (n=1)

The overall median observation time (randomisation to last available assessment) at the cut-off date was 34.4 months (range: 0.1–54.5 months) in the R-chemo arm and 34.8 months (range: 0.0–53.8 months) in the G-chemo arm. The proportion of patients who had been observed for at least 2 years at the clinical cut-off was 87.7% in the R-chemo arm and 91.3% in the G-chemo arm. At the clinical cut-off date, 44.1% of patients in the R-chemo arm and 45.1% of patients in the G-chemo arm had been followed for at least 3 years.

The median duration of post-treatment follow-up at the cutoff date was 9.2 months (range: 0.0–42.3 months) in the R-chemo arm and 9.4 months (range: 0.0–46.9 months) in the G-chemo arm.

During the induction phase, 7.8% patients in the R-chemo arm and 6.2% patients in the G-chemo arm of the FL population were withdrawn from treatment. Most withdrawals were due to AEs and comparable between treatment arms.

During the maintenance phase, 22.0% patients in the R-chemo arm and 19.6% patients in the G-chemo arm of the FL population were withdrawn from treatment. The main reason for withdrawals was progressive disease with a higher proportion of patients in the R-chemo arm (10.6% compared with 6.2% in the G-chemo arm).

Table 24: Reasons for withdrawal by study phase

Reasons for withdrawal, n (%)	G-chemo n=601	R-chemo n=601
Withdrawn from induction phase	**37 (6.2)**	**47 (7.8)**
Adverse event	19 (3.2)	19 (3.2)
Death	3 (0.5)	1 (0.2)
Non-compliance	0	1 (0.2)
Other	2 (0.3)	2 (0.3)
Physician decision	1 (0.2)	5 (0.8)
Progressive disease	5 (0.8)	14 (2.3)
Protocol violation	2 (0.3)	2 (0.3)
Withdrawal by subject	5 (0.8)	3 (0.5)
Withdrawn from maintenance phase	**118 (19.6)**	**132 (22.0)**
Adverse event	51 (8.5)	38 (6.3)
Death	3 (0.5)	4 (0.7)
Lost to follow-up	1 (0.2)	1 (0.2)
Non-compliance	2 (0.3)	0
Other	4 (0.7)	3 (0.5)
Physician decision	15 (2.5)	11 (1.8)
Progressive disease	37 (6.2)	64 (10.6)
Protocol violation	0	1 (0.2)
Withdrawal by subject	5 (0.8)	10 (1.7)
Withdrawn from observation phase	**1 (0.2)**	**0**
Non-compliance	1 (0.2)	0

Withdrawn from follow-up phase	**22 (3.7)**	**34 (5.7)**
Death	18 (3.0)	32 (5.3)
Lost to follow-up	2 (0.3)	1 (0.2)
Withdrawal by subject	2 (0.3)	1 (0.2)
Withdrawn from study	**139 (23.1)**	**183 (30.4)**
Adverse event	1 (0.2)	0
Death	17 (2.8)	14 (2.3)
Lost to follow-up	2 (0.3)	2 (0.3)
Non-compliance	4 (0.7)	2 (0.3)
Other	4 (0.7)	2 (0.3)
Physician decision	7 (1.2)	4 (0.7)
Progressive disease	80 (13.3)	125 (20.8)
Protocol violation	1 (0.2)	4 (0.7)
Withdrawal by subject	23 (3.8)	30 (5.0)

4.5.2 In a table describe the characteristics of the participants at baseline for each of the trials. Provide details of baseline demographics, including age, gender and relevant variables describing disease severity and duration and if appropriate previous treatments and concomitant treatment. Highlight any differences between trial groups.

In the FL population, the treatment arms were in general balanced with respect to demographic factors and baseline disease characteristics (F. Hoffmann-La Roche Ltd, 2016). The median age of patients was 59.0 years (range: 23–88 years); overall, more female than male patients were randomised (53.2% vs. 46.8%). The overall median time from first diagnosis to randomisation was 1.5 months (range: 0.0–168.1 months). The majority of patients had an ECOG performance status of 0-1 (96.8%). The greatest proportion of patients comprised intermediate and high-risk FLIPI (37.2% and 41.8% respectively) and FLIPI-2 groups (50.3% and 40.6%, respectively), and Ann Arbor stage III—IV (>91%). Nearly half (43.8%) of patients had a nodal or extra-nodal mass over 7 cm in diameter. There was extra-nodal involvement in 65.6% of patients.

Table 25: Patient demographics and baseline characteristics (ITT population)

	G-chemo n=601	R-chemo n=601
Mean age, years (SD)	58.2 (11.5)	57.7 (12.2)
Male, n (%)	283 (47.1)	280 (46.6)
Mean height, cm (SD)	168.3 (10.0)	168.4 (10.1)
Mean weight, kg (SD)	76.3 (17.9)	75.2 (17.0)
Mean body surface area, m^2 (SD)	1.86 (0.2)	1.84 (0.2)
Mean BMI, kg/m^2 (SD)	26.8 (5.3)	26.4 (5.9)

Race, n (%)		
Caucasian	487 (81.0)	481 (80.0)
Black or African American	3 (0.5)	1 (0.2)
Asian	100 (16.6)	98 (16.3)
American Indian or Alaska Native	0	1 (0.2)
Native Hawaiian or other Pacific islander	1 (0.2)	0
Multiple	0	3 (0.5)
Other	10 (1.7)	17 (2.8)
Geographic region, n (%)		
Eastern Europe	78 (13.0)	79 (13.1)
Western Europe	294 (48.9)	286 (47.6)
North America	75 (12.5)	77 (12.8)
Asia	92 (15.3)	93 (15.5)
Other	62 (10.3)	66 (11.0)
ECOG PS, n (%)	n=600	n=599
0–1	585 (97.5)	576 (96.2)
2	15 (2.5)	23 (3.8)
Ann Arbor Stage, n (%)	n=598	n=597
I	10 (1.7)	8 (1.3)
II	41 (6.9)	44 (7.4)
III	208 (34.8)	209 (35.0)
IV	339 (56.7)	336 (56.3)
FLIPI no. of adverse factors categories 1, n (%)	n=601	n=601
Low (0,1)	128 (21.3)	125 (20.8)
Intermediate (2)	224 (37.3)	223 (37.1)
High (≥3)	249 (41.4)	253 (42.1)
FLIPI no. of adverse factors categories 2, n (%)	n=579	n=586
Low (0,1)	51 (8.8)	55 (9.4)
Intermediate (2)	296 (51.1)	290 (49.5)
High (≥3)	232 (40.1)	241 (41.1)
Bone marrow involvement at BL, n/patients with data (%)	318/592 (53.7)	295/598 (49.3)
Extranodal involvement, n/patients with data (%)	392/601 (65.2)	396/601 (65.9)
Bulky disease at BL (6 cm threshold), n/patients with data (%)	255/600 (42.5)	271/600 (45.2)
Mean time from diagnosis to randomisation, months (range)	6.25 (0.1–121.6)	7.28 (0.0–168.1)
Chemotherapy regimen, n (%)		
Bendamustine	345 (57.4)	341 (56.7)
CHOP	195 (32.4)	203 (33.8)
CVP	61 (10.1)	57 (9.5)

ECOG, Eastern Cooperative Oncology Group performance score; FLIPI, follicular Lymphoma International Prognostic Index; SD, standard deviation

Differences between PET and non-PET populations

The PET and non-PET ITT populations were comparable with respect to baseline demographic and disease characteristics, although there were some minor differences

between them. Baseline demographic and disease characteristics with at least a 5% difference between the non-PET and PET populations were as follows:

- Patients ≥65 years of age (total: 34.3% vs. 28.2%, respectively)
- Asian patients (total: 19.4% vs. 13.4%, respectively)
- Non-hispanic/Latino patients (total: 92.6% vs. 84.9%, respectively)
- Positive BM involvement (total: 48.9% vs. 54.1%, respectively)
- Extranodal involvement (total: 62.6% vs. 68.6%, respectively).

Within the PET population, baseline demographic and disease characteristics with at least a 5% difference between the R-chemo+R and G-chemo+G arms were as follows:

- Patients ≥60 years of age (40.9% vs.46.5% respectively)
- Ann Arbor Stage IV at diagnosis (54.9% vs. 60.1%, respectively)
- Positive BM involvement (49.3% vs. 59.0%, respectively)
- Negative BM involvement (48.6% vs. 40.7%, respectively).

Patient-reported outcomes: baseline values

Mean baseline scores for each of the individual FACT-Lym questionnaire subscales, and of composite FACT-G, TOI and Total scores, as well as of EQ-5D-3L Utility scales were similar between R-chemo and G-chemo treatment arms. Both arms exhibited some impairment in the functioning and lymphoma symptom subscales as noted by mean scores of between 5 and 15 points lower than the maximum possible depending on the subscale.

Table 26: Baseline scores for patient-reported outcome questionnaires (ITT population)

FACT-LYM Scale, mean (SD)	G-chemo n=601	R-chemo n=601
Physical well-being subscale	23.1 (4.9)	23.4 (4.8)
Functional well-being subscale	18.8 (6.0)	18.7 (6.2)
Emotional well-being subscale	17.9 (4.1)	17.6 (4.2)
Social/family well-being subscale	23.3 (4.8)	22.8 (4.9)
Lymphoma subscale	45.5 (9.3)	45.0 (9.4)
Trial outcome index	86.9 (18.1)	86.6 (18.2)
FACT-G total score	82.9 (14.5)	82.4 (14.9)
FACT-Lym total score	128.4 (22.2)	127.4 (22.4)

FACT-Lym, Functional Assessment of Cancer Therapy-lymphoma; -G, -general

4.6 Quality assessment of the relevant randomised controlled trials

[Provide a quality assessment for each RCT listed in section 4.2.]

Critical appraisal of the included RCT was performed using the format provided in the NICE submission template which adhered to the Centre for Reviews and Dissemination (CRD), University of York guidance (Centre for Reviews and Dissemination). A summary is presented in below:

Table 27: Quality assessment of the identified RCT

Study Question	Grade (Yes/No/ Not Clear/N/A)
	GALLIUM (NCT00545688)
Was randomisation carried out appropriately?	Yes
Was the concealment of treatment allocation adequate?	n/a (open-label study)
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	n/a (open-label study)
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No
Did the analysis include an intent-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes

4.7 Clinical effectiveness results of the relevant randomised controlled trials

In accordance with the proposed indication, the data reported in the clinical effectiveness section is that from the subgroup of patients with FL within the ITT population (the majority of patients in the ITT population had FL; 1202/1401 [85.8%]).

The data discussed in this section will be taken from the primary analysis (clinical cut-off 31st January 2016) (F. Hoffmann-La Roche Ltd, 2016), although data (where available) from the updated analysis will also be presented (clinical cut-off 16th September 2016).

Primary endpoint

Investigator-assessed PFS

The primary endpoint was met in GALLIUM as G-chemo+G treatment resulted in a statistically significant increase in investigator-assessed PFS compared with R-chemo+R.

At the time of the analysis, 24.0% of FL patients in the R-chemo+R arm and 16.8% of FL patients in the G-chemo arm had experienced a PFS event as assessed by the investigator since randomisation. The majority of patients had disease progression as the PFS event (130 patients in the R-chemo+R arm, and 80 patients in the G-chemo+G arm).

G-chemo+G therapy significantly reduced the risk of experiencing a PFS event by 34% compared with R-chemo+R treatment (stratified HR 0.66, 95% CI: 0.51, 0.85; p=0.0012) (F. Hoffmann-La Roche Ltd, 2016).

Table 28: Investigator-assessed PFS, FL patients (FL ITT population), stratified analysis

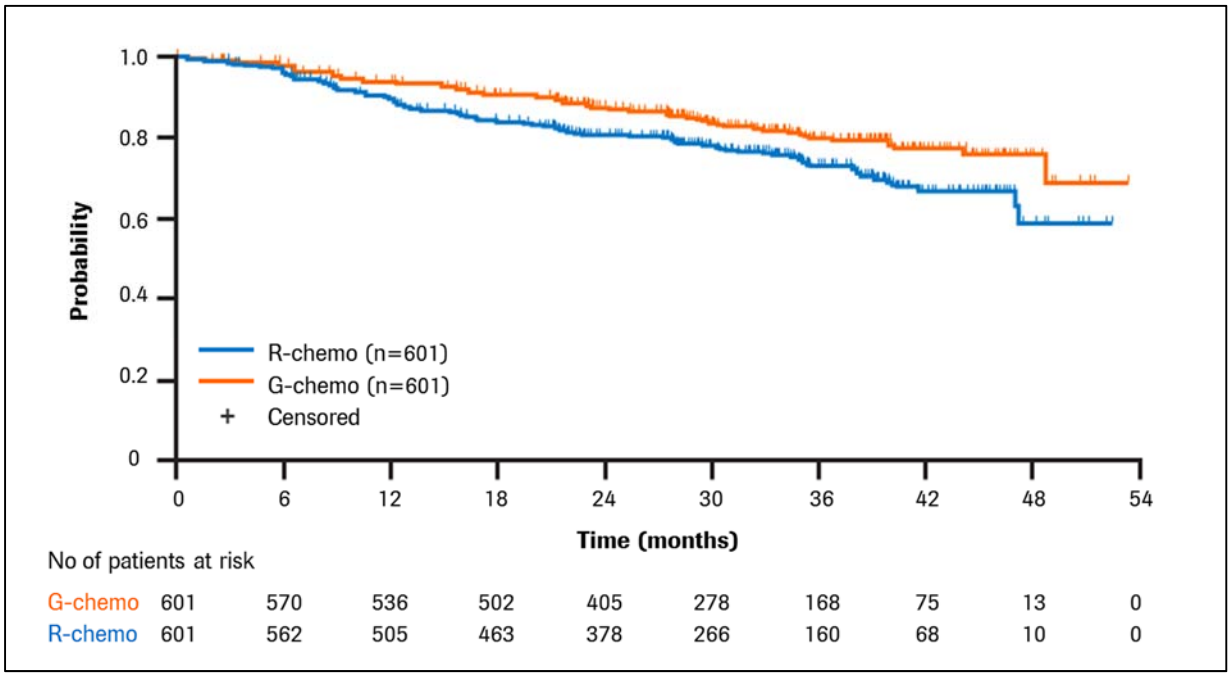
	G-chemo+G n=601	R-chemo+R n=601
Patients with event, n (%)	101 (16.8)	144 (24.0)
Median PFS, months (95% CI)	NE (NE)	NE (47.1, NE)
Hazard ratio (95% CI)	0.66 (0.51, 0.85)	
p value*	0.0012	

*log-rank test
NE, not estimated

The Kaplan–Meier (KM) estimated median PFS times were not reached for either arm. On the basis of KM estimates, 73.3% (95% CI: 68.8, 77.2) of patients in the R-chemo+R arm and 80.0% (95% CI: 75.9, 83.6) of patients in the G-chemo+G arm were progression-free at 3 years. KM estimates are not considered to be reliable beyond the time point when too few patients are at risk (i.e., at least 20% (Pocock et al., 2002)). After 3 years, 160 patients (26.6%) in the R-chemo+R arm, and 168 patients (28.0%) in the G-chemo+G arm were at risk of a PFS event.

A KM plot of investigator-assessed PFS in the FL population is shown below. The KM curves begin to separate in favour of the G-chemo+G arm around 4 months post-randomisation and remain separated thereafter.

Figure 10: KM plot of investigator-assessed PFS, FL patients (FL ITT population)



This result is consistent with that seen in the updated analysis, in which the risk of having a PFS event by investigator assessment was decreased by 32% for patients in the G-chemo+G arm compared with the R-chemo+R (HR 0.68; 95% CI: 0.54, 0.87).

Secondary endpoints

IRC-assessed PFS (FL ITT population)

PFS was also assessed by an IRC; this analysis was consistent with the investigator-assessed findings.

At the time of the analysis, more patients in the R-chemo+R arm experienced a PFS event than in the G-chemo+G arm (20.8% vs 15.5%). Disease progression was recorded for 106 patients in the R-chemo arm and 69 patients in the G-chemo arm (17.6% vs. 11.5%). There were 19 deaths in the R-chemo arm, and 24 deaths in the G-chemo arm before IRC-assessed progression.

The risk of experiencing IRC-assessed disease progression or death was reduced by 29% for patients receiving G-chemo compared to those receiving R-chemo (stratified HR 0.71 [95% CI: 0.54, 0.93]; p=0.0138, stratified log-rank test).

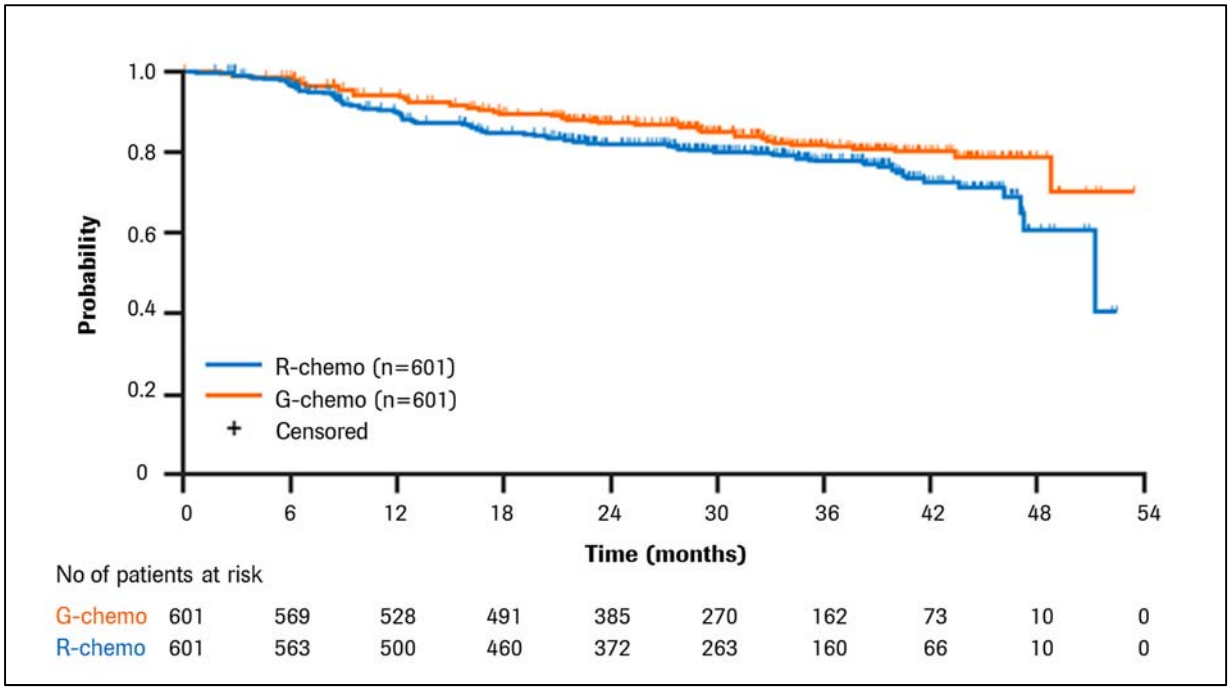
Table 29: IRC-assessed PFS, FL patients (FL ITT population), stratified analysis

	G-chemo+G n=601	R-chemo+R n=601
Patients with event, n (%)	93 (15.5)	125 (20.8)
Median PFS, months (95% CI)	NE (48.7, NE)	51.2 (47.1, NE)
Hazard ratio (95% CI)	0.71 (0.54, 0.93)	
p value*	0.0138	

*log-rank test
NE, not estimated

On the basis of KM estimates, 77.9% (95% CI: 73.8, 81.4) of patients in the R-chemo+R arm and 81.9% (95% CI: 77.9, 85.2) of patients in the G-chemo+G arm were progression-free at 3 years. After 3 years, 160 patients (26.6%) in the R-chemo arm, and 162 patients (27.0%) in the G-chemo were at risk of a PFS event.

Figure 11: KM plot of IRC-assessed PFS, FL patients (FL ITT population)



Sensitivity analyses of PFS (FL ITT population)

The robustness of the PFS results (by investigator or IRC review) was assessed by performing a series of pre-specified sensitivity analyses applying alternative censoring rules or specific analysis criteria. The stratified hazard ratios with 95% confidence intervals and the p-values from the stratified log-rank tests for the primary analyses of PFS based on the Investigators' and IRC's assessments are summarised together with the results from the sensitivity analyses below.

Overall, the results of the panel of pre-specified sensitivity analyses on both Investigator and IRC-assessed PFS in the FL ITT population were consistent, demonstrating the robustness of the primary endpoint. With one exception, the sensitivity analyses produced hazard ratios less than 1.0, indicating better outcomes for patients treated with G-chemo+G than for patients treated with R-chemo+R, consistent with results of the primary analysis. The one exception was a sensitivity analysis assessing the impact of loss to follow-up in which the two treatment arms were treated differently (patients who withdrew prior to disease progression were considered to have progressed at the next scheduled disease assessment in the G-chemo+G arm, but were censored at the last disease assessment in the R-chemo arm). This sensitivity analysis did not produce a hazard ratio <1.0 using the IRC data. However, these findings may be explained by the different way in which the two treatment arms were treated in the analysis (considered a “worst-case” scenario), and the findings are not thought to reflect clinical reality.

Table 30: Summary of sensitivity analyses for PFS (INV-assessed and IRC-Assessed, FL ITT Population)

Analysis	Censoring rules/specific analysis criteria	Stratified HR (95% CI) (p value) (No. of events G-chemo+G vs R-chemo+R)	
		INV-assessed PFS	IRC-assessed PFS
Unstratified log-rank test		0.66 (0.51, 0.85) (p=0.0013) 101 vs. 144	0.71 (0.55, 0.93) (p=0.0131) 93 vs. 125
Re-randomisation test	Re-run randomisation algorithm a large number (10,000, based on IVRS stratification) of times to assess the sensitivity of the stratified log-rank test to the dynamic randomisation procedure	p=0.0011	p=0.0153
Impact of loss to follow-up ("worst case analysis)	Assigns event outcomes to patients who withdrew prior to disease progression in the G-chemo+G arm at the next scheduled disease assessment date and censored outcomes to patients in the R-chemo+R arm at the last disease assessment date	0.90 (0.72, 1.14) (p=0.40) 139 vs. 144	1.17 (0.92, 1.48) (p=0.20) 153 vs. 125
Impact of missed tumour assessment	If patients missed or had an incomplete response assessment prior to the date of the clinical data cutoff or prior to PD, they were counted as having progressed on the day after their last complete response assessment	0.77 (0.63, 0.94) (p=0.0104) 169 vs. 207	0.83 (0.67, 1.03) (p=0.09) 161 vs. 185
Impact of NALT	Censoring at the initiation of non-protocol-specified anti-lymphoma therapy prior to disease progression, to assess potential confounding of the treatment effect estimates by subsequent therapy	0.64 (0.49, 0.83) (p=0.0008) 95 vs. 137	0.72 (0.54, 0.95) (p=0.0183) 87 vs. 116
Impact of treatment discontinuation	Patients who discontinued the study treatment for other reasons than disease progression or death were counted as having progressed at the time of discontinuation (event was date of last dose for early treatment discontinuations)	0.78 (0.64, 0.95) (p=0.0135) 181 vs. 221	0.83 (0.67, 1.01) (p=0.06) 175 vs. 206
Impact of late death	Patients who died more than 6 months after their last response assessment and showed no sign of progression were censored at the last available response assessment	0.64 (0.49, 0.83) (p=0.0007) 95 vs. 140	0.72 (0.55, 0.95) (p=0.0200) 88 vs. 118

Source: (F. Hoffmann-La Roche Ltd, 2016)
CI, confidence interval; HR, hazard ratio; INV, investigator; IRC, Independent Review Committee; ivrs, Interactive Voice Response System; NALT, new anti-lymphoma treatment; PD, progressive disease; PFS, progression-free survival

Investigator-assessed PFS (overall ITT population)

The investigator-assessed PFS data for the FL ITT population is consistent with that seen in the overall ITT population (N=1401)

G-chemo+G therapy significantly reduced the risk of experiencing a PFS event by 32% compared with R-chemo+R treatment (stratified HR 0.68, 95% CI: 0.54, 0.85; p=0.0009).

Table 31: Investigator-assessed PFS, (overall ITT population), stratified analysis

	G-chemo+G n=702	R-chemo+R n=699
Patients with event, n (%)	122 (17.4)	171 (24.5)
Median PFS, months (95% CI)	NE (48.7, NE)	NE (47.1, NE)
Hazard ratio (95% CI)	0.68 (0.54, 0.85)	
p value*	0.0009	

*log-rank test

End-of-induction response with and without PET

Without PET

Based on the investigator assessment, 86.9% (95% CI: 83.9, 89.5) of patients with FL in the R-chemo arm and 88.5% (95% CI: 85.7, 91.0) patients with FL in the G-chemo arm achieved a CR or PR at the end-of-induction (EOI). The difference in ORR was 1.7% (p=0.33, CMH test).

In the R-chemo arm, the CR rate was 23.8% (95% CI: 20.4, 27.4) and the PR rate was 63.1% (95% CI: 59.1, 66.9); whereas, in the G-chemo arm, the CR rate was 19.5% (95% CI: 16.4, 22.9) and the PR rate was 69.1% (95% CI: 65.2, 72.7). The difference in CR rate was -4.3% (p=0.07, CMH test).

The proportion of patients who were non-responders was 13.1% in the R-chemo arm compared with 11.5% in the G-chemo arm. Patients classified as non-responders included patients with SD (1.3% in R-chemo arm vs. 0.5% in G-chemo arm), PD (4.0% in R-chemo arm vs. 2.3% in G-chemo arm), unable to evaluate (3.5% in R-chemo vs. 4.0%, in G-chemo arm), and missing response information (4.3% in R-chemo arm vs. 4.7% in G-chemo arm).

With PET

PET scan results (assessed according to Cheson et al. 2007) were available for 298 patients in the R-chemo arm, and 297 patients in the G-chemo arm. In contrast to CR rates without

PET, the CR rates with PET tended to favour the G-chemo arm with a 5.6% difference in CR rates (56.7% in the R-chemo arm vs. 62.3% in the G-chemo arm). Overall response rates were similar in the two treatment arms, with a difference of 4.3% in favour of the G-chemo arm (81.5% for R-chemo vs. 85.9% for G-chemo).

The results of IRC-assessed CR rate with PET were consistent with the findings of investigator-assessed CR rate with PET. However, the CR rate at the EOI was higher in the G-chemo arm than the R-chemo arm, with a difference of 11.7% in favour of the G-chemo arm (59.7% in the R-chemo arm vs. 71.4% in the G-chemo arm).

A summary of investigator and IRC-assessed EOI response without and with PET is provided below.

Table 32: Investigator and IRC-assessed EOI response with and without PET (FL ITT Population)

	Investigator-assessed		IRC-assessed	
	G-chemo	R-chemo	G-chemo	R-chemo
Without PET				
Overall response, n (%) 95% CI	532/601 (88.5) (85.7, 91.0)	522/601 (86.9) (83.9, 89.5)	548/601 (91.2) (88.6, 93.3)	529/601 (88.0) (85.2, 90.5)
Δ; p value	1.7%; p=0.33		3.2%; p=0.07	
Complete response, n (%) 95% CI	117/601 (19.5) (16.4, 22.9)	143/601 (23.8) (20.4, 27.4)	171/601 (28.5) (24.9, 32.2)	159/601 (26.5) (23.0, 30.2)
Δ; p value	−4.3%; p=0.07		2.0%; p=0.50	
With PET				
Overall response, n (%) 95% CI	255/297 (85.9) (81.4, 89.6)	243/298 (81.5) (76.7, 85.8)	263/297 (88.6) (84.4, 91.9)	254/298 (85.2) (80.7, 89.1)
Δ; p value	4.3%; p=0.19		3.3%; p=0.30	
Complete response, (%) 95% CI	185/297 (62.3) (56.5, 67.8)	169/298 (56.7) (50.9, 62.4)	212/297 (71.4) (65.9, 76.5)	178/298 (59.7) (53.9, 65.4)
Δ; p value	5.6%; p=0.28		11.7%; p=0.0056	

PET, positron emission tomography

Overall survival

At the clinical cutoff date (31st January, 2016), a total of 81 randomised patients had died: 46/601 patients (7.7%) in the R-chemo+R arm and 35/601 patients (5.8%) in the G-chemo+G arm; less than 20% of patients had been followed for survival for more than 4

years, hence the data can be considered still immature at this time (stratified HR 0.75 [95% CI:0.49, 1.17], stratified log-rank p=0.21).

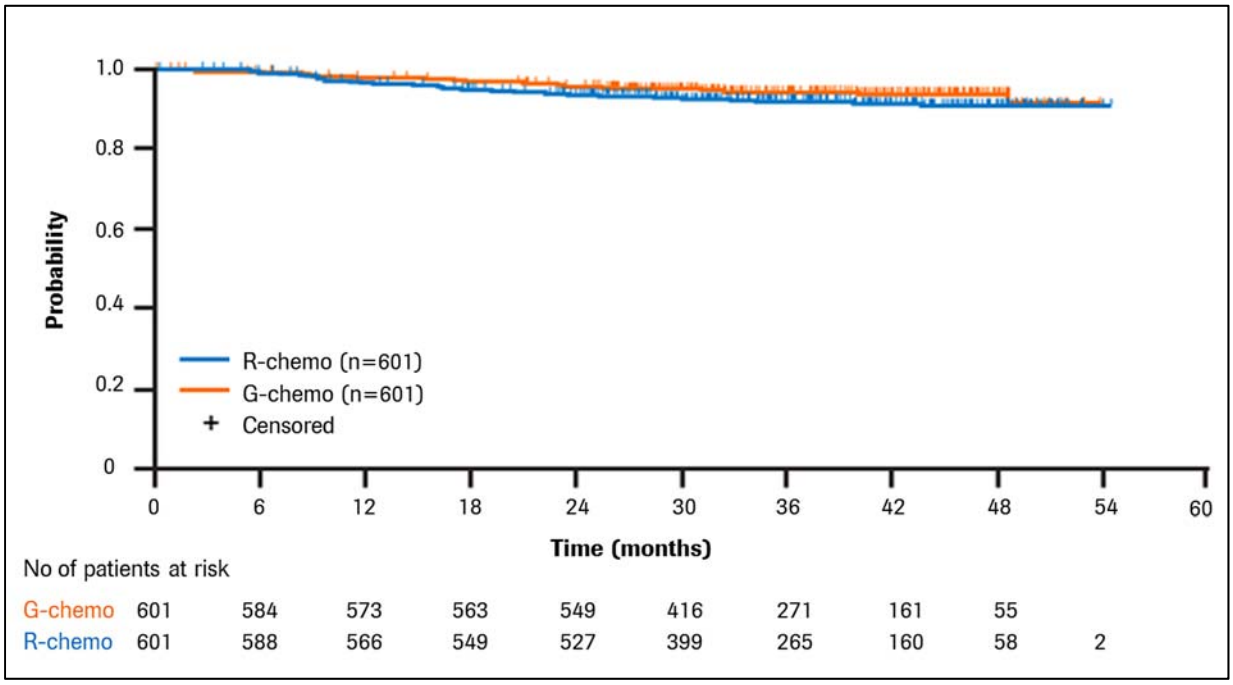
Table 33: Overall survival, (FL ITT population), stratified analysis

	G-chemo+G n=601	R-chemo+R n=601
Patients with event, n (%)	35 (5.8)	46 (7.7)
Median time to event, months (95% CI)	NE (NE, NE)	NE (NE, NE)
Hazard ratio (95% CI)	0.75 (0.49, 1.17)	
p value*	0.21	

*log-rank

On the basis of KM estimates, the estimated probabilities of being alive at 3 years were 92.1% (95% CI: 89.5, 94.1) in the R-chemo+R arm and 94.0% (95% CI: 91.6, 95.7) in the G-chemo+G arm. Based on visual inspection, the KM plot for OS showed a separation of curves favouring the G-chemo+G treatment arm. Median overall survival time had not been reached in either of the treatment arms.

Figure 12: KM plot of overall survival, FL patients (FL ITT population)



This result is consistent with that seen in the updated analysis; HR 0.82 (95% CI: 0.54, 1.22)

Event-free survival

In the R-chemo+R arm, 26.5% of patients had experienced an EFS event (PD, death or start of NALT) compared to 18.6% of patients in the G-chemo+G arm. Disease progression accounted for most of the earliest contributing EFS events, and the difference in PD events

between the two arms accounted for the major difference in EFS events between the arms. The proportion of patients who had died (2.7%, 32/1202) or had NALT (3.2%, 39/1202) as the earliest contributing EFS event was similar in each treatment arm.

Compared to the R-chemo+R arm, patients randomized to the G-chemo+G arm were significantly less likely to experience disease progression, death or start NALT (stratified HR 0.65 [95% CI: 0.51, 0.83]; p=0.0006, stratified log-rank test).

Table 34: Event-free survival, (FL ITT population), stratified analysis

	G-chemo+G n=601	R-chemo+R n=601
Patients with event, n (%)	112 (18.6)	159 (26.5)
Median time to event, months (95% CI)	NE (47.1, NE)	NE (NE, NE)
Hazard ratio (95% CI)	0.65 (0.51, 0.83)	
p value*	0.0006	

*log-rank

Disease-free survival

Disease-free survival was assessed in patients with a response of CR (as assessed by the investigator) any time prior to NALT. A total of 281/601 patients (46.8%) in the R-chemo+R arm, and 298/601 patients (49.6%) in the G-chemo+G arm experienced a CR before commencement of NALT. By the clinical cut-off date, 11.7% of patients in the R-chemo arm, and 9.1% of patients in the G-chemo+G arm experienced disease progression or death. Treatment with G-chemo+G reduced the risk of progression or death in patients with a CR by 19% compared to patients with a CR who received treatment with R-chemo+R arm (stratified HR 0.81 [95% CI: 0.48; 1.35], stratified log-rank test).

Table 35: Disease-free survival, FL patients with CR (FL ITT population), stratified analysis

	G-chemo+G n=298	R-chemo+R n=281
Patients with event, n (%)	27 (9.1)	33 (11.7)
Median time to event, months (95% CI)	NE (NE, NE)	NE (NE, NE)
Hazard ratio (95% CI)	0.81 (0.48, 1.35)	

CI, confidence interval; NE, not estimated

Duration of response

Duration of response was assessed in patients with a CR or PR (by investigator assessment) prior to NALT. This included 94% of patients in the R-chemo+R arm, and 95% of patients in the G-chemo+G arm. Of these responders, 21.9% of patients in the

R-chemo+R arm, and 15.4% of patients in the G-chemo+G arm subsequently had disease progression or death. The HR for responders progressing or dying was 0.66 (95% CI: 0.50, 0.87) in favour of the G-chemo+G arm. Treatment with G-chemo+G reduced the risk of progression or death in patients with a CR or PR by 34% compared to patients with a CR or PR who received treatment with R-chemo+R. Note that the group of responders in the G-chemo+G arm may be different from the responders in the R-chemo+R arm.

Table 36: Duration of response in patients with CR/PR, FL patients (FL ITT population), stratified analysis

	G-chemo+G n=571	R-chemo+R n=567
Patients with event, n (%)	88 (15.4)	124 (21.9)
Median time to event, months (95% CI)	NE (NE, NE)	NE (44.5, NE)
Hazard ratio (95% CI)	0.66 (0.50, 0.87)	

CI, confidence interval; NE, not estimated

Time to new anti-lymphoma treatment

At the time of the clinical cutoff, 18.5% of patients in the R-chemo+R arm, and 13.3% of patients in the G-chemo arm had started a NALT or died from any cause. Compared to the R-chemo+R arm, patients randomised to the G-chemo+G arm were less likely to start a NALT or die from any cause (stratified HR 0.68 [95% CI: 0.51, 0.91];p=0.0094, stratified log-rank test).

New anti-lymphoma treatments were received by 14.8% of patients in the R-chemo+R arm, and 9.7% of patients in the G-chemo+G arm. New anti-lymphoma treatments which were received by >5 patients included:

- R-CHOP (17 patients in the R-chemo+R arm, and 7 patients in the G-chemo+G arm)
- Radiotherapy (10 patients in the R-chemo+R arm, and 6 patients in the G-chemo+G arm)
- MabThera monotherapy (14 patients in the R-chemo+R arm and 11 patients in the G-chemo+G arm)
- Bendamustine combined with MabThera (6 patients in the R-chemo+R arm and 8 patients in the G-chemo+G arm)
- Transplantation (10 patients in the R-chemo+R arm, and 7 patients in the G-chemo+G arm).

Very few patients received a NALT before a PFS event which included 7/601 patients (1.2%) in the R-chemo+R arm, and 6/601 patients (1.0%) in the G-chemo+G arm.

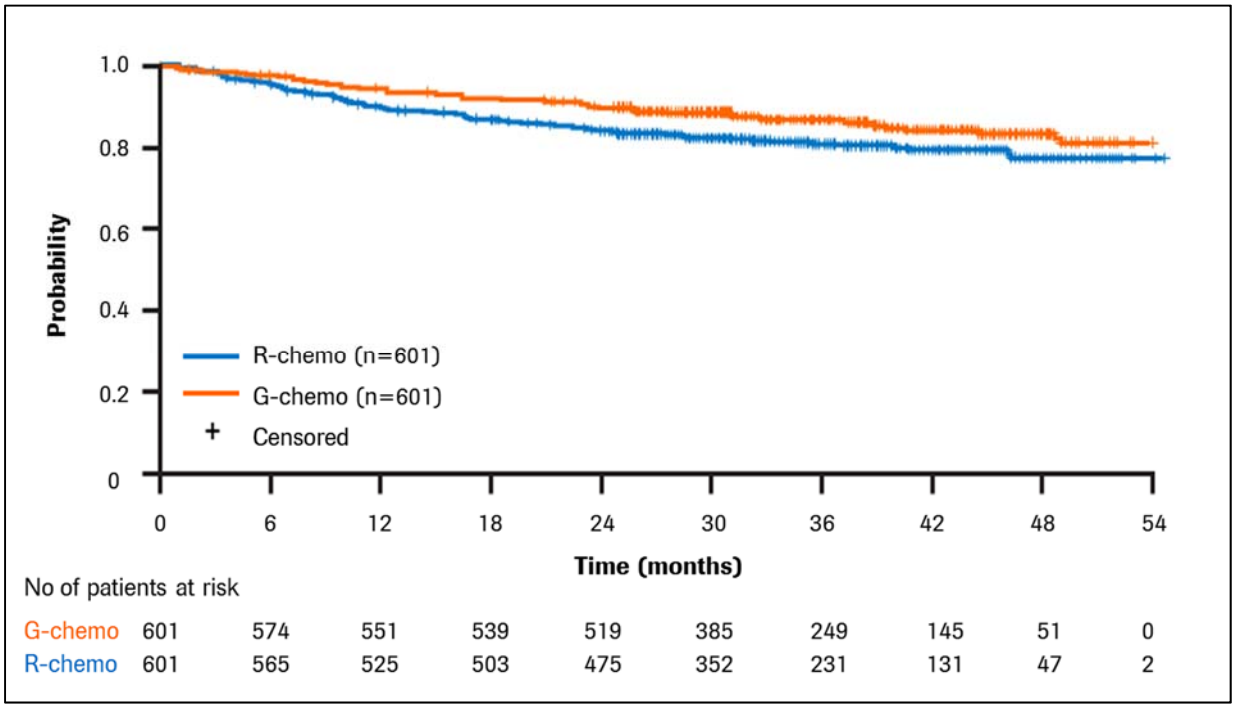
Table 37: Time to new anti-lymphoma therapy (FL ITT population), stratified analysis

	G-chemo+G n=601	R-chemo+R n=601
Patients with event, n (%)	80 (13.3)	111 (18.5)
Median time to event, months (95% CI)	NE (NE, NE)	NE (NE, NE)
Hazard ratio (95% CI)	0.68 (0.51, 0.91)	
p value*	0.0094	

*log-rank
CI, confidence interval; NE, not estimated

On the basis of KM estimates, the estimated probabilities at 3 years to be NALT-free or surviving were 81.2% (95% CI: 77.6, 84.2) in the R-chemo+R arm and 87.1% (95% CI: 84.0, 89.6) in the G-chemo+G arm. Visually, the KM curves begin to separate in favour of the G-chemo+G arm around 4 months post-randomisation and remain separated thereafter.

Figure 13: KM plot of time to NALT (FL ITT population)



Patient reported-outcomes

The proportions of patients randomised to each treatment arm who completed all scales on the FACT-Lym and EQ-5D questionnaires were generally balanced between treatment arms, suggesting that differences in attrition rate between the two arms can be ruled out as a potential confounder in the analysis results.

Table 38: FACT-Lym and EQ-5D questionnaire compliance by visit

% compliance	G-chemo+G n=601	R-chemo+R n=601
FACT-Lym		
Baseline	92.5	91.5
Completion of induction	82.9	84.6
Completion of maintenance	78.2	76.5
Follow-up month 36	67.4	72.0
EQ-5D		
Baseline	93.0	92.8
Completion of induction	83.4	85.1
Completion of maintenance	77.7	76.0
Follow-up month 36	66.5	70.5

Mean baseline scores for each of the individual FACT-Lym questionnaire subscales, and of composite FACT-G, TOI and Total scores, as well as of EQ-5D-3L utility scales were similar between R-chemo+R and G-chemo+G treatment arms. Both arms exhibited some impairment in the functioning and lymphoma symptom subscales as noted by mean scores of between 5 and 15 points lower than the maximum possible depending on the subscale.

Table 39: Mean baseline FACT-Lym questionnaire scale scores

FACT-Lym scale, mean (SD)	G-chemo+G n=601	R-chemo+R n=601
Physical Well-being	23.14 (4.85)	23.36 (4.77)
Functional Well-being	18.76 (5.98)	18.66 (6.19)
Emotional Well-being	17.87 (4.13)	17.64 (4.19)
Social/Family Well-being	23.28 (4.77)	22.84 (4.92)
Lymphoma Subscale	45.54 (9.29)	45.01 (9.37)
Trial Outcome Index	86.94 (18.05)	86.61 (18.16)
FACT-G Total	82.92 (14.52)	82.35 (14.87)
FACT-Lym Total	128.42 (22.16)	127.40 (22.43)

Note: Max score: PWB: 28, FWB: 28, EWB: 24, SFWB: 28, Lyms: 60, TOI: 116, G: 108, Total:168

There were no notable differences between the treatment arms in any of the FACT-Lym questionnaire subscales or EQ-5D-3L scales over time during the induction and maintenance treatment periods, and follow-up, as evidenced by modest (<5%) between arm differences in the mean changes from baseline scores in FACT-Lym subscales, TOI and Total score, and EQ-5D-3L Utility scales.

Equal proportions of patients in the G-chemo+G and R-Chemo+R arms had improvement in their FACT-Lym questionnaire scores during treatment and throughout maintenance and follow-up as defined by a ≥3 point increase from baseline in the Lymphoma subscale, a ≥6 point increase from baseline in the FACT Lym TOI and a ≥7 point increase from baseline in

the FACT Lym Total score. By the Month 2 maintenance assessment, although the number of patients available for assessment was lower, approximately 50% of the patients still in the treatment arms were reporting clinically meaningful improvement.

Table 40: Summary of meaningful Improvement in FACT-Lym

FACT-Lym Subscale (definition of meaningful improvement), n (%)	G-chemo+G n=601	R-chemo+R n=601
Lymphoma subscale (≥ 3 point increase)		
Cycle 3, Day 1 (Induction treatment)	229 (45.1)	217 (40.8)
End of Induction visit	233 (47.0)	238 (47.6)
Maintenance visit Month 2	233 (57.4)	212 (56.5)
Maintenance visit Month 12	227 (53.7)	216 (56.1)
Maintenance Completion visit	218 (56.2)	205 (55.0)
FACT TOI (≥ 6 point increase)		
Cycle 3, Day 1 (Induction treatment)	162 (31.7)	163 (30.5)
End of Induction visit	189 (38.0)	203 (40.0)
Maintenance visit Month 2	192 (47.1)	182 (48.3)
Maintenance visit Month 12	202 (47.6)	190 (49.1)
Maintenance Completion visit	191 (49.1)	174 (46.4)
FACT Total (≥ 7 point increase)		
Cycle 3, Day 1 (Induction treatment)	173 (33.9)	179 (33.5)
End of Induction visit	197 (39.6)	206 (40.6)
Maintenance visit Month 2	191 (46.8)	180 (47.7)
Maintenance visit Month 12	197 (46.5)	188 (48.5)
Maintenance Completion visit	191 (49.1)	171 (45.5)

Note: Percentages are calculated on the number of patients who completed the questionnaire at each visit.

Further information on the patient-reported outcomes in GALLIUM, including EQ-5D data will be discussed in section 5.4 of this submission.

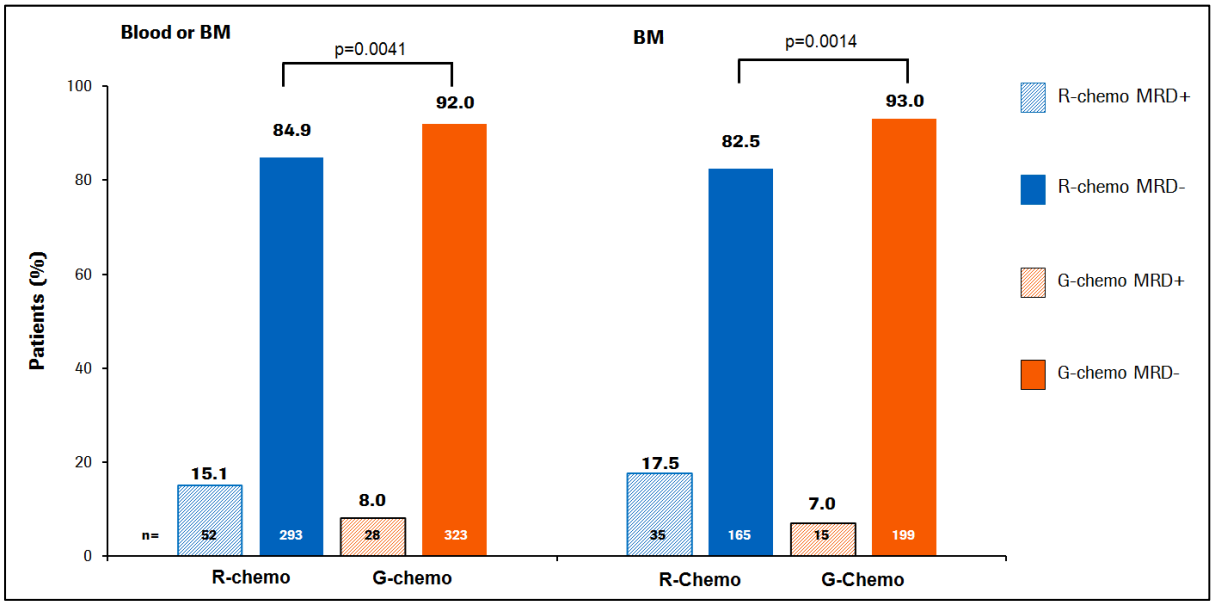
Exploratory endpoints

MRD analysis

Of the 1202 FL patients enrolled in GALLIUM, 1138 provided consent for MRD analyses. Baseline peripheral blood (PB) or bone marrow (BM) samples were available for 1101 patients; a clonal marker was detected in 968 (88%) of these patients and 815 (74%) had an real-time quantitative-polymerase chain reaction (RQ-PCR) assay fulfilling sensitivity criteria (Pott C et al., 2016). Baseline characteristics were comparable between patients with a detectable clonal marker to those without, with the exception of higher-stage disease (61% vs 34% for Ann Arbor stage IV), reflecting an increased BM involvement.

Among the 696 patients with an available PB or BM sample at EOI, MRD response was significantly higher in the G-chemo+G arm than the R-chemo+R arm (92% vs 85%; p=0.0041).

Figure 14: MRD status by compartment arm at end-of-induction



Source: (Pott C et al., 2016)
BM, bone marrow; MI, mid-induction; MRD, minimal residual disease; PB, peripheral blood

MRD clearance occurred early during treatment: at mid-induction, 94% of patients in the G-chemo+G arm achieved MRD-negative status in PB compared with 89% in the R-chemo+R arm (p=0.013).

Table 41: MRD status by treatment arm at MI in PB

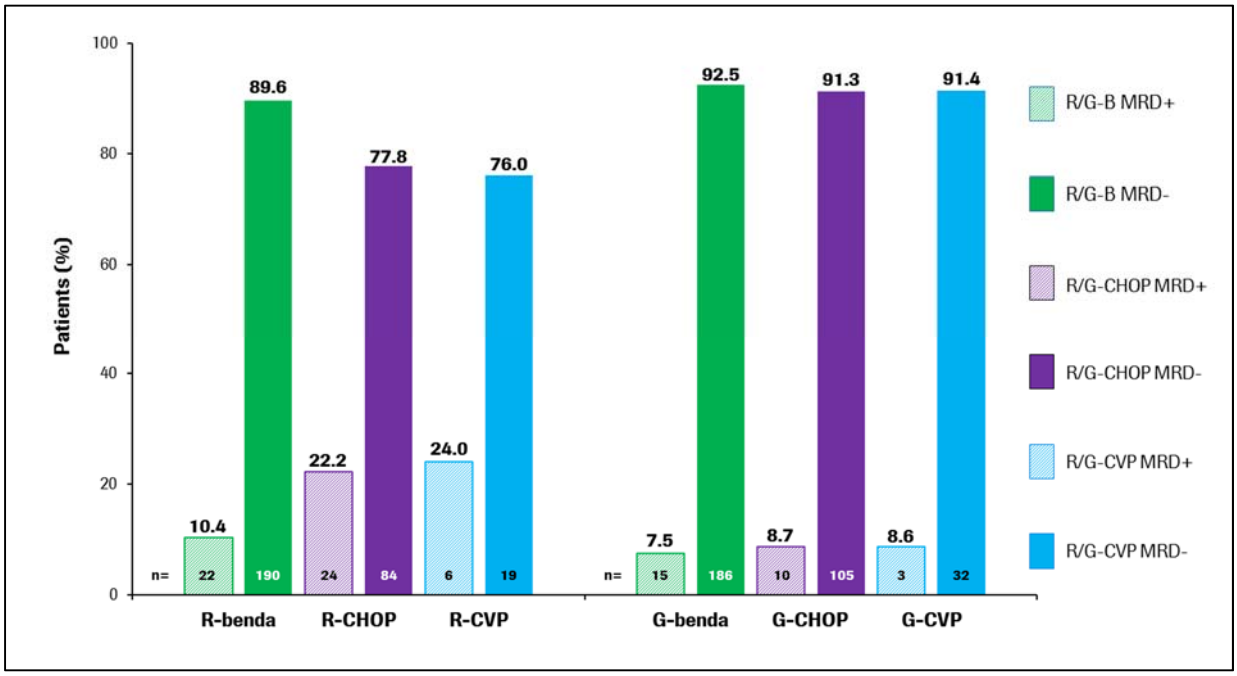
MRD status at MI in PB, n (%)	G-chemo n=348	R-chemo n=342
MRD positive	20 (5.7)	38 (11.1)
MRD negative	328 (94.3)	304 (88.9)
p-value	0.013	

Source: (Pott C et al., 2016)
BM, bone marrow; MI, mid-induction; MRD, minimal residual disease; PB, peripheral blood

The anti-lymphoma activity of G-chemo induction was confirmed by analysing quantitative MRD data in PB at MI: all 20 (100%) patients who remained MRD-positive at MI in the G-chemo arm had low-level MRD (below the limit of quantification) compared with 24/38 (63%) patients in the R-chemo arm.

The chemotherapy backbone in the R-chemo arm affected MRD status in PB and BM at EOI (MRD-negativity rates 89.6%, 77.8% and 76.0% after R-bendamustine, R-CHOP and R-CVP, respectively); however, no such effect was seen in the G-chemo arm where MRD response rates at EOI were high and similar with all three chemo regimens.

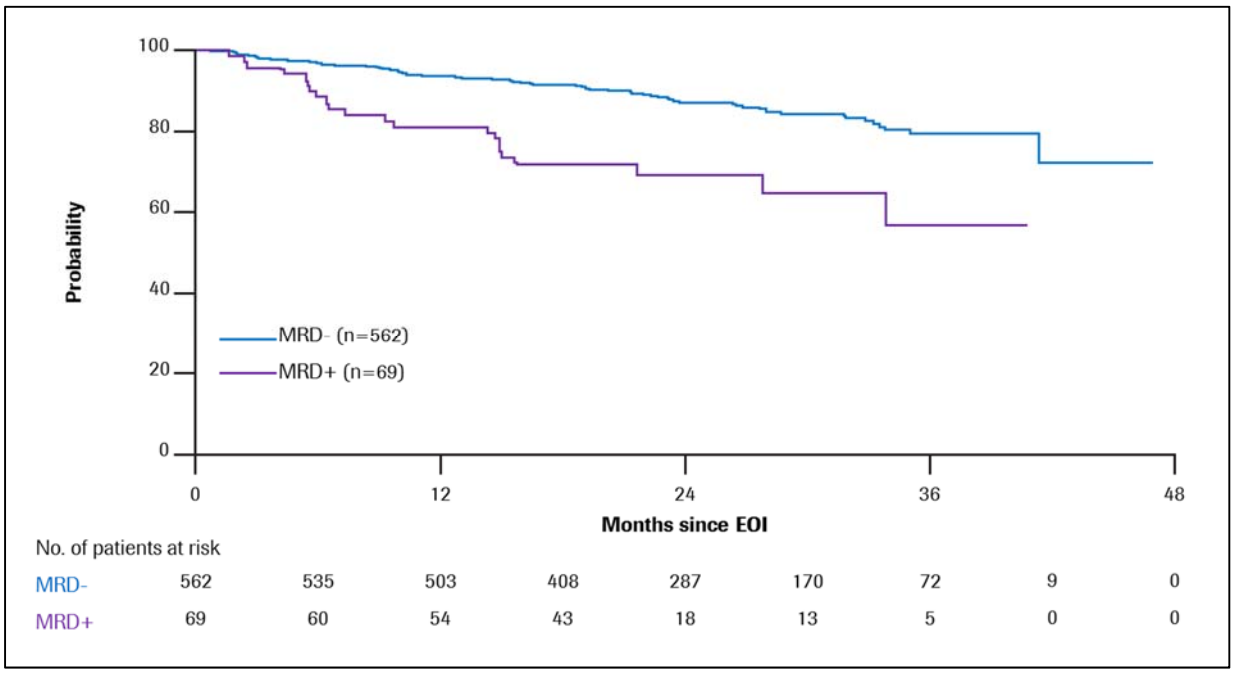
Table 42: MRD status by chemotherapy regimen and treatment arm at EOI in PB and BM



Source: (Pott C et al., 2016)

Achievement of MRD negativity at EOI in PB/BM for patients with CR/PR at EOI was associated with longer subsequent PFS, with a hazard ratio of 0.35 (95% CI, 0.22, 0.56; p<0.0001) and comparable effects in both treatment arms.

Figure 15: PFS from EOI by MRD status (patients receiving maintenance)



Source: (Pott C et al., 2016)

End-of-maintenance response

For medical interest, a non-pre-specified end of maintenance response (EOMR) analysis was conducted. EOMR was defined as the first response assessment that occurred after the last dose of maintenance treatment. The population included all patients randomised other than those who had not yet reached the end of the maintenance assessment. Patients with disease progression or death at any time or with missing response assessments at the end of maintenance were considered non-responders.

Based on the Investigator assessment, 712/1058 patients with FL (67.3%) achieved a CR or PR at the end of maintenance phase: 341 patients (64.0% [95% CI: 59.7, 68.1]) in the R-chemo+R arm and 371 patients (70.7% [95% CI: 66.6, 74.5]) in the G-chemo+G arm, an absolute difference of 6.7% (95% CI: 1.0, 12.4; p-value=0.0197, CMH test) in favour of G-chemo+G at this time point. Similarly, an absolute difference in CR rate of 2.5% (95% CI: -3.5, 8.4; p-value=0.39, CMH test) was found in favour of G-chemo at this time point.

Table 43: Overall end-of-maintenance response (without PET) (FL ITT population)

	G-chemo+G n=525	**R-chemo+R n=533**
Overall response (CR, PR)		
n (%)	371 (70.7)	341 (64.0)
(95% CI)	(66.6, 74.5)	(59.74 68.1)
Difference G-chemo+G vs R-chemo+R	6.69	
(95% CI)	(0.95, 12.43)	
p-value*	0.0197	
Complete response		
n (%)	205 (39.0)	195 (36.6)
(95% CI)	(34.9, 43.4)	(32.5, 40.8)
Difference G-chemo+G vs R-chemo+R	2.46	
(95% CI)	(-3.48, 8.41)	
p-value*	0.3871	
Partial response		
n (%)	166 (31.6)	146 (27.4)
(95% CI)	(27.7, 35.8)	(23.7, 31.4)
Difference G-chemo+G vs R-chemo+R	4.23	
(95% CI)	(-1.36, 9.82)	
p-value*	0.1389	
Stable disease		
n (%)	0 (0)	1 (0.2)
(95% CI)	(0.0, 0.7)	(0.0, 1.0)
Difference G-chemo+G vs R-chemo+R	-0.19	
(95% CI)	(-0.65, 0.28)	
p-value*	0.2733	
Progressive disease		
n (%)	41 (7.8)	70 (13.1)
(95% CI)	(5.7, 10.5)	(10.4, 16.3)

Difference G-chemo+G vs R-chemo+R (95% CI) p-value*	-5.32 (-9.1, -1.6) 0.0039

*based on stratified CMH

4.8 Subgroup analysis

Subgroup analyses of investigator-assessed PFS (FL ITT population)

The potential impact of baseline demographics, prognostic factors, and stratification factors on the treatment effect was assessed. Hazard ratios for PFS with 95% confidence intervals (G-chemo+G vs. R-chemo+R) for pre-specified patient subgroups are shown on the forest plots below. With the exception of FL FLIPI low risk (HR 1.17 [95% CI: 0.63, 2.19]; based on 253 patients), the observed hazard ratios were below 1.00 and ranged from 0.40–0.86 for subgroups including at least 10% of patients. Overall, the results of the PFS subgroup analyses are consistent with the primary analysis of PFS in the FL population.

The majority of investigators chose bendamustine (57%) and <10% of investigators chose CVP as the backbone chemotherapy regimen for patients at their site. Regardless of chemotherapy regimen, PFS was better in patients randomised to G-chemo+G. The observed hazard ratios by chemotherapy subgroup were as follows; CHOP (n=398): HR 0.77 (95% CI: 0.50, 1.20), CVP (n=118): HR 0.63 (95% CI: 0.32, 1.21), and bendamustine (n=686): HR 0.61 (95% CI: 0.43, 0.86). Subgroup analyses for the different chemotherapy regimens should be interpreted with caution because the trial was not designed to compare the efficacy of chemotherapy. The induction regimen was chosen on a per centre basis for patients with FL. Accordingly, there could be differences in patient populations treated with the different regimens.

The potential impact of baseline demographics, prognostic factors and stratification factors on the treatment effect as assessed by the IRC was also analysed. The results of the IRC-assessed PFS subgroup analyses are consistent with the overall analysis of IRC-assessed PFS in the FL ITT population, and with the investigator-assessed PFS subgroup analysis

Figure 16: Subgroup analyses of investigator-assessed PFS by stratification factors (FL ITT population)

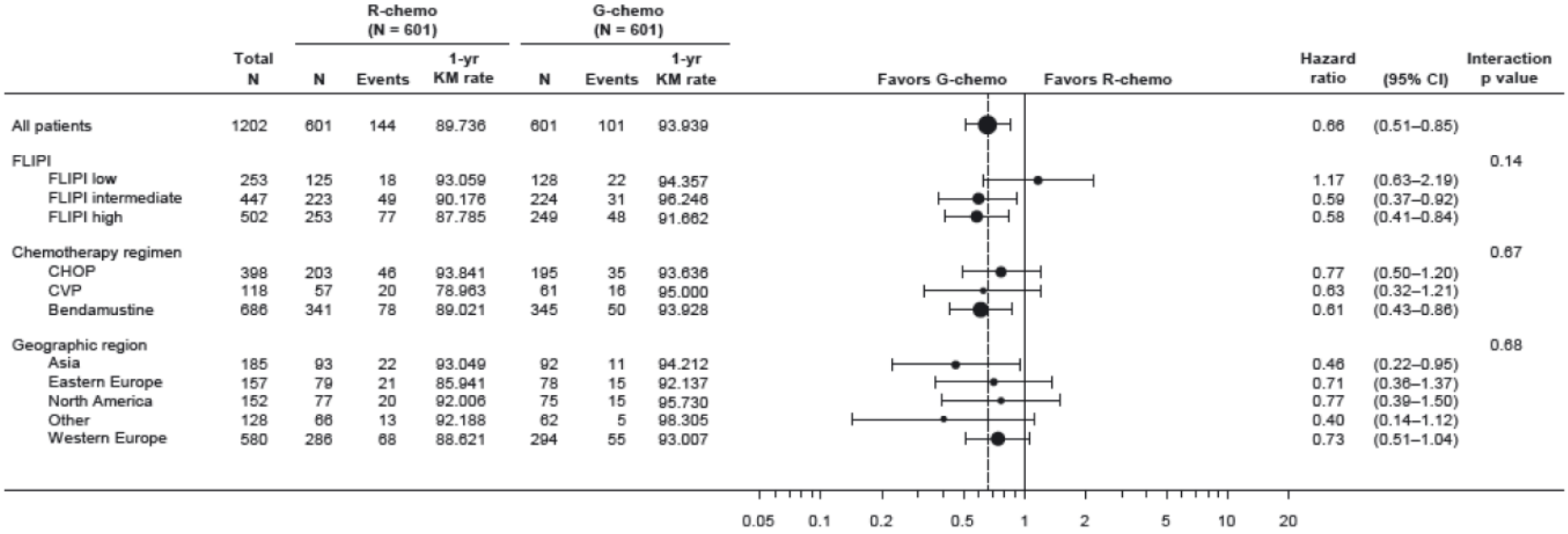


Figure 17: Subgroup analyses of investigator-assessed PFS by age (FL ITT population)

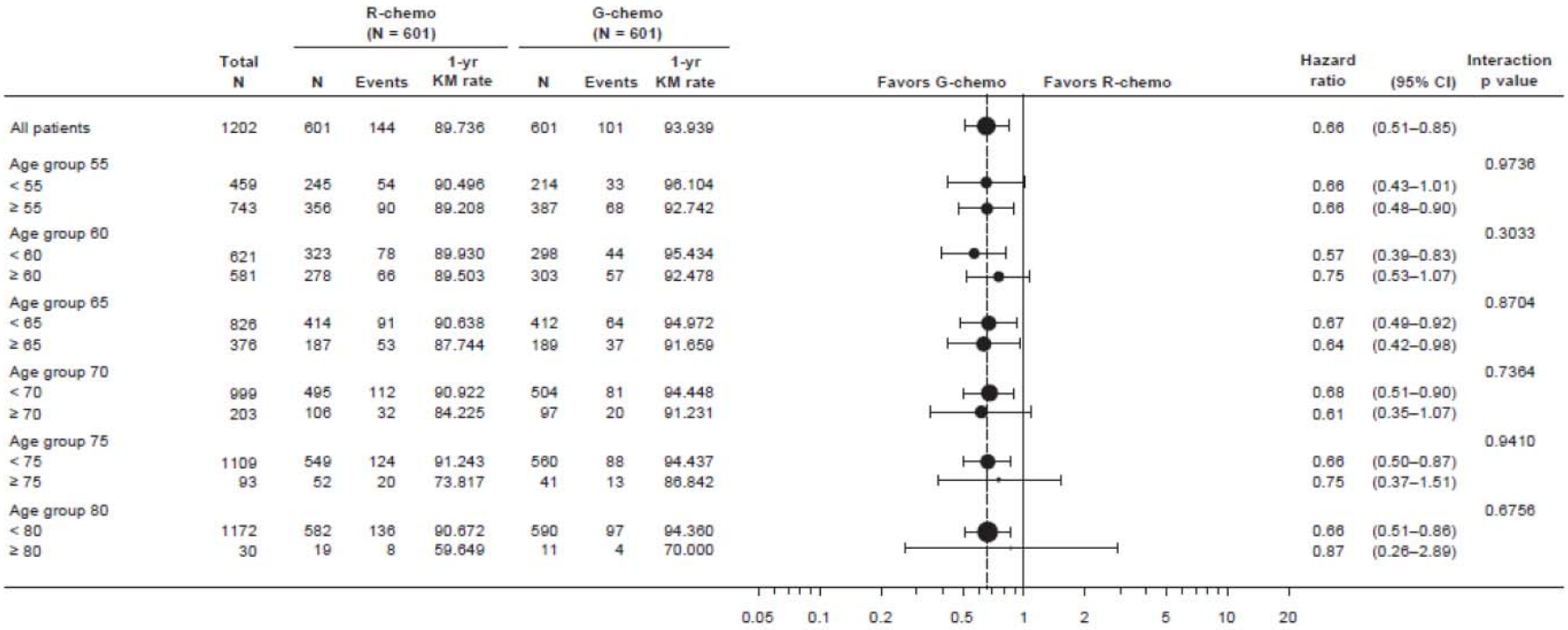
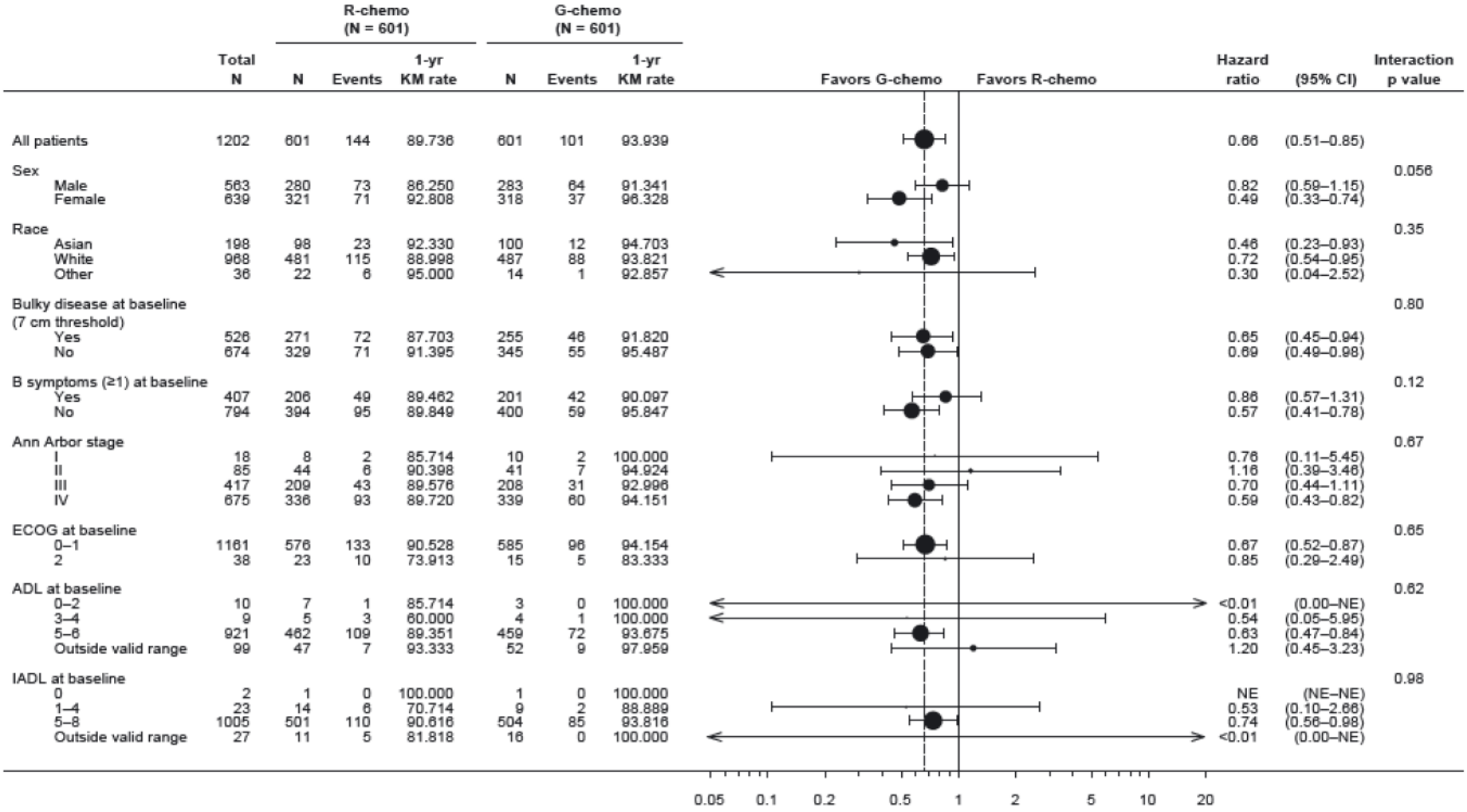


Figure 18: Subgroup analyses of investigator-assessed PFS by baseline demographics and disease characteristics (FL ITT population)



4.9 Meta-analysis

GALLIUM was the only randomised clinical study identified in the SLR to be relevant to the decision problem therefore a meta-analysis is not feasible.

4.10 Indirect and mixed treatment comparisons

No indirect and mixed treatment comparisons were conducted as the GALLIUM study addressed all comparators highlighted in the decision problem.

4.11 Non-randomised and non-controlled evidence

Provide details of the non-randomised and non-controlled studies that provide additional evidence to supplement RCT data. Provide a list of the relevant studies and summarise the methodology, statistical analyses, participant flow and quality assessment for each. Briefly summarise the results of the non-randomised and non-controlled studies.

The efficacy and safety of Gazyvaro in combination with chemotherapy (CHOP or bendamustine) as induction, followed by Gazyvaro monotherapy as maintenance in previously-untreated patients with FL has been investigated in an open-label, non-randomised Phase Ib study (GAUDI, NCT 00825149) (Grigg et al., 2016).

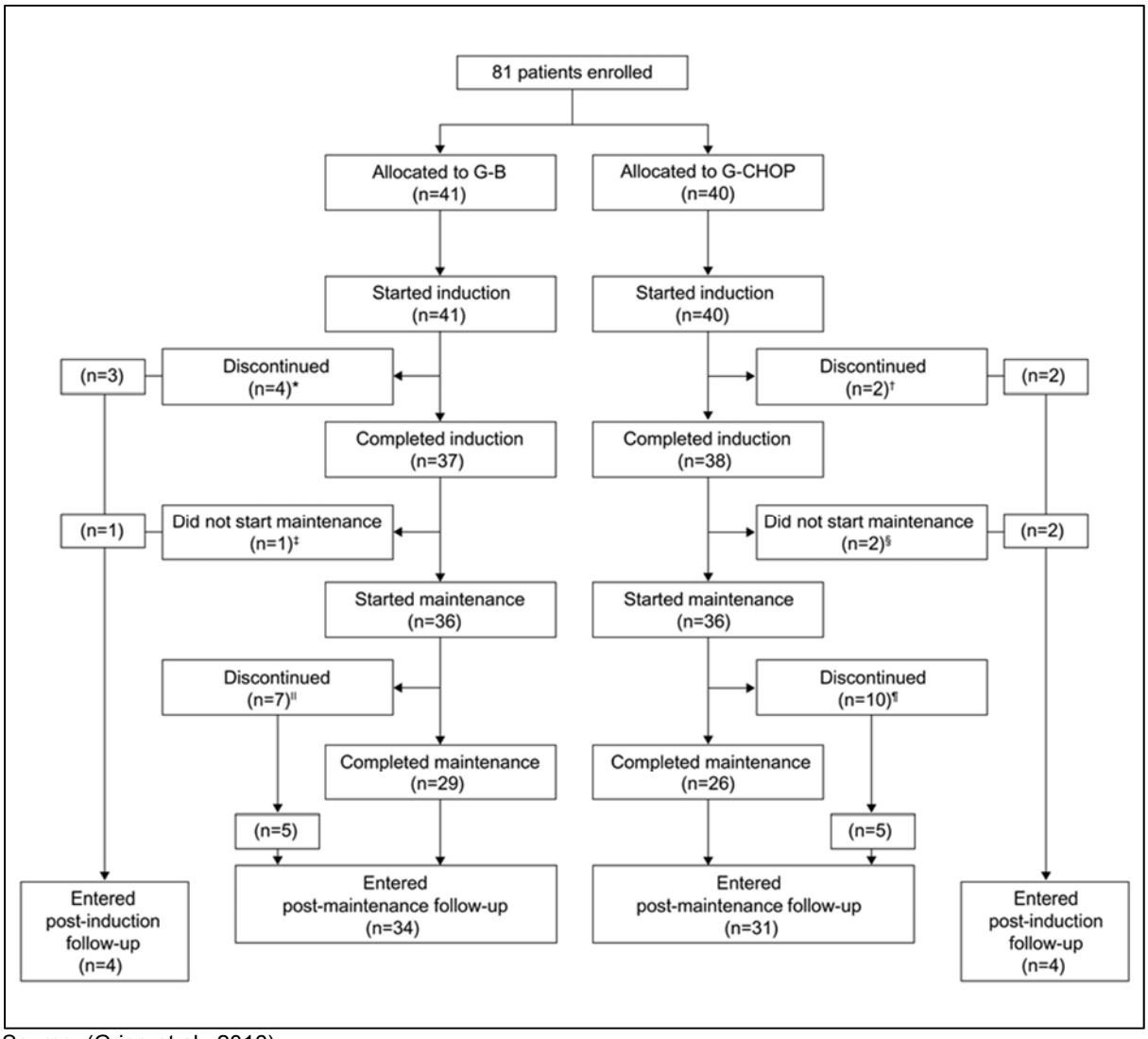
Eighty-one patients were enrolled; 41 were allocated to the G-benda group and 40 to the G-CHOP group. The majority of patients (91%) were Ann Arbor stage III–IV, had an intermediate/high FLIPI score (82%) and had extra-nodal involvement (67%); 43% had bulky disease.

Assignment to chemotherapy regimen was decided on a per centre basis before enrolment. Patients received Gazyvaro (1000 mg intravenously [iv], days 1 and 8 of cycle 1, and day 1 of subsequent cycles) plus bendamustine (4–6 cycles at 4-week intervals: 90 mg/m^2 iv on days 2 and 3 of cycle 1, and days 1 and 2 of subsequent cycles) or CHOP (6-8 cycles at 3-week intervals: cyclophosphamide, 750 mg/m^2 iv day 1; doxorubicin, 50 mg/m^2 iv day 1; vincristine, 1.4 mg/m^2 capped at 2 mg iv day 1; prednisone, 100 mg orally days 1–5). Patients with a CR or PR at the EOI were eligible for maintenance with Gazyvaro monotherapy (1000 mg iv) starting 12 weeks after the last chemoimmunotherapy dose and administered every 3 months for 2 years or until PD.

Eighty patients were planned for the safety evaluation. All patients who received ≥1 dose of G-chemotherapy were eligible for the safety and efficacy analyses. For the efficacy evaluation, response rates and 95% Pearson-Clopper CIs were estimated. PFS was assessed using Kaplan-Meier methodology.

The treatment allocation and study flow is summarised below.

Figure 19: Patient disposition in Phase Ib non-randomised study, GAUDI



Source: (Grigg et al., 2016)
*Reasons for discontinuation from G-B induction therapy: insufficient therapeutic response (n=2), administrative/other (n=1), and withdrawal of consent (n=1; this patient did not enter post-induction follow-up).
†Reasons for discontinuation from G-CHOP induction therapy: adverse event (AE)/intercurrent illness (n=1) and administrative/other (n=1). ‡Reasons patients did not start G-maintenance treatment (G-B group): AE/intercurrent illness (n=1).
§Reasons patients did not start: G-maintenance treatment (G-CHOP group): administrative/other (n=2).
‖Reasons for withdrawal from maintenance treatment (G-B group): AE/intercurrent illness (n=5) and insufficient therapeutic response (n=2).
¶Reasons for withdrawal from maintenance treatment (G-CHOP group): AE/intercurrent illness (n=4), insufficient therapeutic response (n=3), administrative/other (n=2), and death (n=1).

Safety – induction phase

All patients experienced at least 1 AE during the induction phase, with 64% (G-benda, 51%; G-CHOP, 78%) experiencing grade 3/4 AEs. IRRs were the most common AE (58% of patients); the majority occurred during cycle 1 and were grade ≤2 in intensity. The most common grade 3/4 haematologic AE was neutropenia, occurring in 36% of patients (G-benda, 29%; G-CHOP, 43%) during induction, although febrile neutropenia was rare. Grade 3/4 non-haematologic AEs overall were uncommon; Grade 3/4 infections occurred in 13 patients (16%), predominantly in the context of neutropenia (9 patients). *P. jirovecii* pneumonia was reported in 1 patient.

Safety – maintenance phase

Overall, 27 of 72 eligible patients experienced Grade 3-5 AEs during maintenance. Nine patients withdrew from G treatment due to an AE, 5 in the G-benda group (due to giardiasis with anaemia, neutropenic infection, flare-up of Crohn's disease, nasopharyngitis, and neutropenia in 1 patient each) and 4 in the G-CHOP group (3 due to infection and 1 due to peripheral sensory neuropathy).

The most common class of non-haematologic AEs was infections, with 11 patients (G-benda, 6; G-CHOP, 5) experiencing a variety of grade 3 infections and 1 patient in the G-benda group experiencing a Grade 4 neutropenic infection. No further cases of *P. jirovecii* pneumonia were reported during maintenance.

Eight patients experienced haematologic AEs during maintenance, all in the G-benda group. Six patients (8%) experienced Grade 3/4 neutropenia (n=5) or febrile neutropenia (n=1), noted 81-91 days after the last dose of Gazyvaro.

Safety – follow-up phase

No serious adverse events (SAEs) were observed in the 8 patients who entered follow-up directly post-induction. Three patients experienced SAEs during post-maintenance follow-up. In the G-B group, 1 patient had lower abdominal pain (Grade 3); in the G-CHOP group, 1 patient each had an abnormal liver function test (Grade 4) and dyspnoea (Grade 3).

Efficacy

The ORR was 94% at the EOI; the estimated PFS rate at 36 months was 87%. At the final analysis, 17 events defining progression/death had occurred in 81 patients: one event (progression) occurred during induction, 6 during maintenance (5 progression and 1 death,

including 1 patient in the G-CHOP group with transformation to diffuse large B-cell lymphoma), and 10 after maintenance.

Table 44: Efficacy parameter summary, GAUDI

Variable	G-benda n=41	G-CHOP n=40	Total N=81
ORR, % (95% CI)	93 (80.1, 98.5)	95 (83.1, 99.4)	94 (86.2, 98.0)
CR at end of induction, % (95% CI)	37 (22.1, 53.1)	35 (20.6, 51.7)	36 (25.4, 47.2)
CR at 30 months, % (95% CI)	63 (46.0, 78.2)	58 (40.8, 74.5)	61 (NA, NA)
PFS at 36 months, % (95% CI)	90 (0.80, 0.99)	84 (0.72, 0.96)	87 (0.79, 0.94)
Progression/death (n)	6	11	17
Deaths due to PD (n)	1	2	3

Source: (Grigg et al., 2016)

The results from this Phase Ib study demonstrate that induction therapy with G-benda or G-CHOP, followed by Gazyvaro maintenance, is associated with tolerable safety and promising efficacy.

Table 45: Quality assessment of the non-randomised controlled trials

Study Question	Grade (Yes/No/ Not Clear/N/A)
	GAUDI (NCT00825149)
Was randomisation carried out appropriately?	N/A
Was the concealment of treatment allocation adequate?	n/a (open-label study)
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	n/a (open-label study)
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No
Did the analysis include an intent-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes

4.12 Adverse reactions

4.12.1 Evidence from comparative RCTs and regulatory summaries is preferred, but findings from non comparative trials may sometimes be relevant. For example, post

marketing surveillance data may demonstrate that the technology shows a relative lack of adverse reactions commonly associated with the comparator, or that the occurrence of adverse reactions is not statistically significantly different to those associated with other treatments.

4.12.2 In a table, summarise adverse reactions reported in the studies listed in section 4.2. For each intervention group, give the number with the adverse reaction and the frequency, the number in the group, and the percentage with the reaction. Then present the relative risk and risk difference and associated 95% confidence intervals for each adverse reaction.

The data presented in this section are from the FL safety analysis population (i.e. patients with FL who received any amount of study drug [Gazyvaro, MabThera, or chemotherapy: CHOP, CVP, or bendamustine]) from the primary analysis of the GALLIUM study (clinical cut-off 31st January 2016) (F. Hoffmann-La Roche Ltd, 2016). Safety analyses were found to be comparable to the overall iNHL population; a comparison of the overall safety results between these populations is provided in below.

There was a numerically higher rate of deaths (for any reason, including progressive disease) in the overall population compared to the FL population. The incidence of Grade 3-5 AEs were comparable in the FL and overall populations

Table 46: Comparison of safety analyses in the FL and the overall safety populations

n, (%)	FL		Overall	
	G-chemo+G n=595	R-chemo+R n=597	G-chemo+G n=698	R-chemo+R n=692
No. of patients with at least one AE (any Grade)	592 (99.5)	587 (98.3)	695 (99.6)	682 (98.6)
Total no. of events	10,311	9,343	12,364	10,702
Total no. of deaths	35 (5.9)	46 (7.7)	50 (7.2)	63 (9.1)
No. of patients with at least one AE				
AE with fatal outcome	24 (4.0)	20 (3.4)	36 (5.2)	26 (3.8)
Grade 3–5 AE	444 (74.6)	405 (67.8)	528 (75.6)	479 (69.2)
SAE	274 (46.1)	238 (39.9)	340 (48.7)	286 (41.3)
SAE leading to treatment withdrawal	44 (7.4)	36 (6.0)	54 (7.7)	50 (7.2)
SAE leading to dose reduction	12 (2.0)	10 (1.7)	14 (2.0)	13 (1.9)
SAE leading to dose interruption	83 (13.9)	45 (7.5)	109 (15.6)	55 (7.9)
Related SAE	152 (25.5)	122 (20.4)	193 (27.7)	149 (21.5)
AE leading to treatment withdrawal	97 (16.3)	85 (14.2)	125 (17.9)	104 (15.0)
AE leading to dose reduction	107 (18.0)	95 (15.9)	133 (19.1)	109 (15.8)
AE leading to dose interruption	395 (66.4)	338 (56.6)	474 (67.9)	402 (58.1)
Related AE	564 (94.8)	547 (91.6)	663 (95.0)	634 (91.6)
Related AE leading to treatment withdrawal	75 (12.6)	65 (10.9)	100 (14.3)	80 (11.6)
Related AE leading to dose reduction	103 (17.3)	89 (14.9)	129 (18.5)	101 (14.6)

Related AE leading to dose interruption	349 (58.7)	296 (49.6)	422 (60.5)	350 (50.6)

*any treatment
AE, adverse event; FL, follicular lymphoma; G-chemo+G, Gazyvaro + chemotherapy followed by Gazyvaro maintenance; R-chemo+R, MabThera + chemotherapy followed by MabThera maintenance

Extent of exposure

A total of 1192 patients with FL received any amount of study drug during the induction phase (597 patients in the R-chemo arm, and 595 patients in the G-chemo arm), and are included in the FL safety population.

During induction, most patients received all planned doses of Gazyvaro or MabThera. The median duration of treatment with MabThera and Gazyvaro during induction was the same in the two arms (25.1 weeks).

As summarised below, 526 patients in the R-chemo+R arm received R-maintenance treatment, and 540 patients in the G-chemo+G arm received G-maintenance treatment. At the time of the clinical cut-off date, 114 patients with FL were still ongoing with maintenance treatment (54 in the R-chemo arm and 60 in the G-chemo arm). The median duration of treatment with MabThera and Gazyvaro during maintenance was the same in the two arms (92 weeks).

Table 47: Summary of exposure (induction phase) (safety population)

	Induction												
	G-chemo+G							R-chemo+R					
	G n=595	R* n=3	B n=338	C n=254	H n=193	P n=255	V/O n=254	R n=597	B n=338	C n=259	H n=203	P n=259	V/O n=259
Median treatment duration, wks (range)	25.1 (3.3–35.3)	14.1 (4.1–24.1)	24.3 (3.9–31.4)	24.3 (3.9–30.0)	24.3 (3.9–30.0)	24.3 (3.9–30.0)	24.3 (3.9–30.0)	25.1 (2.6–32.3)	24.3 (3.9–30.0)	19.3 (2.6–28.1)	19.1 (3.9–30.0)	19.9 (2.4–28.9)	19.3 (2.6–28.1)
Dose intensity, %													
<60%	0.3	33.3	0	0.4	0	0.4	4.3	0	0	0	0.5	0.4	4.6
60-<80%	0	0	3.0	3.9	4.1	2.0	9.4	0	4.1	2.3	2.5	2.3	6.9
80-<90%	0	33.3	6.5	5.1	5.7	3.5	5.5	0.5	6.5	1.9	2.0	3.9	5.0
≥90%	99.7	0	90.5	90.6	90.2	94.1	80.7	99.5	89.3	95.8	95.1	93.4	83.4
Missing	0	33.3	0	0	0	0	0	0	0	0	0	0	0

	Maintenance		
	G n=540	R* n=3	R n=526
Median treatment duration, wks (range)	92.3 (0.0–117.3)	4.1 (0.0–98.6)	92.1 (2.1–117.7)
Dose intensity, %			
<60%	0	0	0
60-<80%	0	0	0
80-<90%	0	0	0.8
≥90%	99.8	33.3	99.2
Missing	0.2	66.7	0

B, bendamustine; C, cyclophosphamide; G, Gazyvaro; H, doxorubicin; P, prednisone; R, MabThera; V/O, vincristine
Treatment duration is the date of the last dose of study medication minus the date of the first dose plus 28 days, or if new anti-leukaemia therapy was started within these 28 days exposure duration is the time interval between first dose and start of new anti-leukaemia therapy minus 1 day.
Dose intensity is the total dose actually received divided by the total planned dose.
*Three patients received MabThera in error

Common adverse events

The incidence of AEs over the entire study period (i.e., induction, maintenance and follow-up) was similar in the two treatment arms; 98.0% had at least one AE in the R-chemo+R arm compared with 99.5% in the G-chemo+G arm. The most frequently affected System Organ Classes were as follows (percentages expressed as R-chemo+R vs. G-chemo+G):

- Gastrointestinal disorders (75.2% vs. 79.3%)
- Infections and infestations (70.0% vs. 77.3%)
- General disorders and administration site conditions (68.8% vs. 74.5%)
- Injury, poisoning and procedural complications (55.1% vs. 63.9%)
- Blood and lymphatic system disorders (52.8% vs. 58.3%).

The five most frequently reported AEs were (percentages expressed as R-chemo+R vs. G-chemo+G):

- Infusion-related reactions (IRRs) (48.9% vs. 59.0%),
- Nausea (46.6% vs. 46.9%)
- Neutropenia (43.6% vs 48.6%)
- Fatigue (36.5% vs. 36.0%)
- Constipation (31.5% vs 35.3%).

AEs that occurred with ≥2% difference in incidence between treatment arms (excluding IRRs) are presented in Table 48 below.

Table 48: Adverse events that occurred with ≥2% difference in incidence rate between treatment arms (excluding IRRs) (safety population)

n, (%)	G-chemo+G n=595	R-chemo+R n=597
Total number of patients	509 (85.5)	504 (84.4)
Total number of AE, n	2266	1983
Blood and lymphatic system disorders		
Number of patients with at least one AE	318 (53.4)	278 (46.6)
Neutropenia	289 (58.6)	260 (43.6)
Thrombocytopenia	62 (10.4)	45 (7.5)
Febrile neutropenia	42 (7.1)	29 (4.9)
Gastrointestinal disorders		
Number of patients with at least one AE	359 (60.3)	345 (57.8)
Nausea	187 (31.4)	214 (35.8)
Constipation	188 (31.6)	173 (29.0)
Diarrhoea	147 (24.7)	127 (21.3)
Dyspepsia	47 (7.9)	29 (4.9)
General disorders and administration site conditions		

Number of patients with at least one AE	15 (2.5)	29 (4.4)
Pain	15 (2.5)	29 (4.4)
Infections and infestations		
Number of patients with at least one AE	150 (25.2)	106 (17.8)
Herpes zoster	59 (9.9)	39 (6.5)
Sinusitis	55 (9.2)	38 (6.4)
Rhinitis	41 (6.9)	26 (4.4)
Pharyngitis	26 (4.4)	13 (2.2)
Metabolism and nutrition disorders		
Number of patients with at least one AE	38 (6.4)	22 (3.7)
Hypokalaemia	38 (6.4)	22 (3.7)
Musculoskeletal and connective tissue disorder		
Number of patients with at least one AE	76 (12.8)	95 (15.9)
Back pain	76 (12.8)	95 (15.9)
Psychiatric disorders		
Number of patients with at least one AE	78 (13.1)	65 (10.9)
Insomnia	78 (13.1)	65 (10.9)
Skin and subcutaneous tissue disorders		
Number of patients with at least one AE	78 (13.1)	64 (10.7)
Alopecia	78 (13.1)	64 (10.7)

Treatment-related adverse events

Related AEs were observed in 91.6% of patients in the R-chemo+R arm and 94.8% of patients in the G-chemo+G arm. Related AEs were most frequently reported in the following System Organ Classes (percentages expressed as R-chemo+R vs. G-chemo+G):

- Gastrointestinal disorders (62.0% vs. 65.2%)
- General disorders and administration site conditions (50.8% vs. 60.8%)
- Injury, poisoning and procedural complications (49.1% vs. 59.2%)
- Blood and lymphatic system disorders (48.2% vs. 54.3%).

Adverse events by severity

The majority of AEs were Grade 1 or 2 in severity in each arm (85.9% in the R-chemo+R arm and 85.0% in the G-chemo+G arm). A total of 1319 AEs in the R-chemo+R arm and 1544 AEs in the G-chemo+G arm were Grade 3–5 in severity.

Table 49: Summary of AEs by highest Grade (safety population)

n, (%)	G-chemo+G n=595	R-chemo+R n=597
Total		
Patients with at least one AE, n (%)	592 (99.5)	585 (98.0)
Total number of AEs, n	10,311	9,341
Grade 1		
Patients with at least one AE, n (%)	15 (2.5)	22 (3.7)

Total number of AEs, n	5,531	5,531
Grade 2		
Patients with at least one AE, n (%)	133 (22.4)	158 (26.5)
Total number of AEs, n	3,236	3,005
Grade 3		
Patients with at least one AE, n (%)	210 (35.3)	216 (36.2)
Total number of AEs, n	1,044	933
Grade 4		
Patients with at least one AE, n (%)	210 (35.3)	169 (28.3)
Total number of AEs, n	474	366
Grade 5		
Patients with at least one AE, n (%)	24 (4.0)	20 (3.4)
Total number of AEs, n	26	20

Multiple occurrences of the same AE in the same individual are counted in the total number of AEs.

The incidence of Grade 3–5 AEs during the entire treatment period was higher in the G-chemo+G arm (74.6%) than in the R-chemo+R arm (67.8%); this was driven by a higher incidence (≥2% higher incidence in G-chemo+G vs. R-chemo+R) of neutropenia, febrile neutropenia, IRRs, and thrombocytopenia. All Grade 3–5 AEs (by preferred term) reported in ≥2% of patients with FL in the study are summarised in Table 50.

Table 50: Grade 3-5 AEs reported in ≥2% of patients with FL in either treatment arm (Safety Population)

n, (%)	G-chemo+G n=595	R-chemo+R n=597
Neutropenia*	261 (43.9)	226 (37.9)
Leukopenia	51 (8.6)	50 (8.4)
Febrile neutropenia*	41 (6.9)	29 (4.9)
Infusion-related reaction*	40 (6.7)	22 (3.7)
Thrombocytopenia*	36 (6.1)	16 (2.7)
Pneumonia	29 (4.9)	26 (4.4)
Anaemia	24 (4.0)	13 (2.2)
Dyspnoea	17 (2.9)	9 (1.5)
Hypertension	14 (2.4)	10 (1.7)

*values for these preferred terms have a ≥2% higher incidence in Grade 3–5 AE in the G-chemo+G arm compared to the R-chemo+R arm.

Serious adverse events

Overall, there was a higher incidence of SAEs in the G-chemo+G arm than in the R-chemo+R arm. A total of 238/597 patients (39.9%) in the R-chemo+R arm experienced 450 SAEs compared with and 274/595 patients (46.1%) in the G-chemo arm, experiencing 590 SAEs.

Table 51: Serious adverse events over the entire study period, occurring in ≥1% patients (safety population)

n, (%)	G-chemo+G n=595	R-chemo+R n=597
Total number of patients with at least one event	274 (46.1)	238 (39.9)
Total number of AE, n	590	450
Blood and lymphatic system disorders		
Number of patients with at least one AE	56 (9.4)	47 (7.9)
Febrile neutopenia	29 (4.9)	19 (3.2)
Neutropenia	22 (3.7)	25 (4.2)
Gastrointestinal disorders		
Number of patients with at least one AE	43 (7.2)	28 (4.7)
Diarrhoea	8 (1.3)	6 (1.0)
Abdominal pain	8 (1.3)	5 (0.8)
Vomiting	3 (0.5)	7 (1.2)
General disorders and administration site conditions		
Number of patients with at least one AE	30 (5.0)	34 (5.7)
Pyrexia	18 (3.0)	17 (2.8)
Infections and infestations		
Number of patients with at least one AE	108 (18.2)	86 (14.4)
Pneumonia	29 (4.9)	25 (4.2)
Herpes zoster	6 (1.0)	8 (1.3)
Urinary tract infection	8 (1.3)	5 (0.8)
Infection	5 (0.8)	7 (1.2)
Lower respiratory tract infection	8 (1.3)	3 (0.5)
Lung infection	5 (0.8)	6 (1.0)
Sepsis	8 (1.3)	2 (0.3)
Bronchitis	6 (1.0)	3 (0.5)
Gastroenteritis	7 (1.2)	1 (0.2)
Injury, poisoning and procedural complications		
Number of patients with at least one AE	41 (6.9)	21 (3.5)
Infusion-related reactions	27 (4.5)	11 (1.8)
Respiratory, thoracic and mediastinal disorders		
Number of patients with at least one AE	33 (5.5)	30 (5.0)
Dyspnoea	6 (1.0)	6 (1.0)
Pulmonary embolism	6 (1.0)	2 (0.3)
Vascular disorders		
Number of patients with at least one AE	12 (2.0)	7 (1.2)
Hypotension	6 (1.0)	0

Adverse events of particular or special interest

The frequency and severity of AE of particular or special interest in GALLIUM was consistent with the known safety profile of Gazyvaro.

n, (%)	G-chemo+G n=595	R-chemo+R n=597
Infusion-related reactions*		
Number of patients with at least one AE	406 (68.2)	349 (58.5)

Number of patients with Grade 3–5 AEs	40 (6.7)	22 (3.7)
Number of patients with serious AEs	33 (5.5)	0 (0.0)
Neutropenia		
Number of patients with at least one AE	301 (50.6)	269 (45.1)
Number of patients with Grade 3–5 AEs	261 (43.9)	226 (37.9)
Number of patients with serious AEs	50 (8.4)	44 (7.4)
Infections		
Number of patients with at least one AE	460 (77.3)	418 (70.0)
Number of patients with Grade 3–5 AEs	118 (19.8)	93 (15.6)
Number of patients with serious AEs	108 (18.2)	86 (14.4)
Tumour lysis syndrome		
Number of patients with at least one AE	6 (1.0)	3 (0.5)
Number of patients with Grade 3–5 AEs	6 (1.0)	3 (0.5)
Number of patients with serious AEs	3 (0.5)	1 (0.2)
Thrombocytopenia		
Number of patients with at least one AE	68 (11.4)	45 (7.5)
Number of patients with Grade 3–5 AEs	36 (6.1)	16 (2.7)
Number of patients with serious AEs	4 (0.7)	1 (0.2)
Acute thrombocytopenia		
Number of patients with at least one AE	7 (1.2)	0 (0.0)
Number of patients with Grade 3–5 AEs	5 (0.8)	0 (0.0)
Number of patients with serious AEs	2 (0.3)	0 (0.0)
Haemorrhagic events		
Number of patients with at least one AE	57 (9.6)	62 (10.4)
Number of patients with Grade 3–5 AEs	5 (0.8)	7 (1.2)
Number of patients with serious AEs	6 (1.0)	5 (0.8)
Gastrointestinal perforation		
Number of patients with at least one AE	4 (0.7)	3 (0.5)
Number of patients with Grade 3–5 AEs	3 (0.5)	0 (0.0)
Number of patients with serious AEs	3 (0.5)	0 (0.0)
Cardiac events		
Number of patients with at least one AE	78 (13.1)	58 (9.7)
Number of patients with Grade 3–5 AEs	22 (3.7)	17 (2.8)
Number of patients with serious AEs	0 (0.0)	0 (0.0)
Second malignancies (6 months after first study drug intake)		
Number of patients with at least one AE	62 (10.4)	42 (7.0)
Number of patients with Grade 3–5 AEs	30 (5.0)	17 (2.8)
Number of patients with serious AEs	35 (5.7)	18 (3.0)
Hepatitis B reactivation		
Number of patients with at least one AE	3 (0.5)	2 (0.3)
Number of patients with Grade 3–5 AEs	0 (0.0)	0 (0.0)
Number of patients with serious AEs	0 (0.0)	0 (0.0)

*Most frequent symptoms of IRRs; nausea (24.2% [G-chemo+G], 19.3% [R-chemo+R]), chills (15.0%, 6.9%), pyrexia (13.6%, 5.5%), vomiting (10.4%, 7.5%), fatigue (6.7%, 6.9%)

Deaths

Up until the clinical cut-off date of 31st January 2016, 46/597 patients (7.7%) in the R-chemo+R arm and 35/595 patients (5.9%) in the G-chemo+G arm had died during the study. Progressive disease was considered by the investigator to be the primary cause of death in

22/597 patients (3.7%) in the R-chemo+R arm and 12/595 patients (2.0%) in the G-chemo+G arm. The frequency of deaths due to adverse events was similar in the two arms (3.4% vs 3.9%, respectively).

Table 52: Summary of deaths (safety population)

n, (%)	G-chemo+G n=595	R-chemo+R n=597
Subject status		
Alive	560 (94.1)	551 (92.3)
Dead	35 (5.9)	46 (7.7)
Cause of death		
Adverse event	23 (3.9)	20 (3.4)
Progressive disease	12 (2.0)	22 (3.7)
Other	0 (0.0)	4 (0.7)

Safety by study phase

A summary of the safety results for the FL population in the entire study is provided below. Events were analysed by the phase in which they started, although events starting in one phase could have continued into subsequent phases of the study. Furthermore, events starting in the maintenance or follow-up phases may have been due to treatment received in an earlier phase. Most events (68.9% overall) and most Grade 3– 5 events (69.7% overall) occurred during induction, the period during which Gazyvaro and MabThera were given concurrently with chemotherapy (CHOP, CVP, or bendamustine).

Table 53: Overview of adverse events: entire study and during each phase (safety population)

	Entire study period		Induction		Maintenance		Follow-up	
n (%)	G-chemo+G n=595	R-chemo+R n=597	G-chemo n=595	R-chemo n=597	G n=548	R n=535	G-chemo+G n=444	R-chemo+R n=451
Total number of pts with at least one AE	592 (99.5)	585 (98.0)	580 (97.5)	577 (96.6)	501 (91.4)	458 (85.6)	130 (29.3)	106 (23.5)
Total number of events, n	10,309	9,341	7,012	6,533	3,002	2,578	295	230
Grade 3–5 AE	444 (74.6)	405 (67.8)	357 (60.0)	336 (56.3)	205 (37.4)	169 (31.6)	56 (12.6)	33 (7.3)
Grade 5 AE	24 (4.0)	20 (3.4)	4 (0.7)	3 (0.5)	10 (1.8)	10 (1.9)	10 (2.3)	7 (1.6)
Serious AE	274 (46.1)	238 (39.9)	166 (27.9)	144 (24.1)	134 (24.5)	110 (20.6)	47 (10.6)	34 (7.5)
AE leading to withdrawal*	97 (16.3)	85 (14.2)	47 (7.9)	49 (8.2)	51 (9.3)	36 (6.7)	2 (0.5)	0 (0.0)
AE of Particular interest (Grade 3–5)								
IRR	40 (6.7)	22 (3.7)	39 (6.6)	21 (3.5)	3 (0.5)	1 (0.2)	0 (0.0)	0 (0.0)
Neutropenia	261 (43.9)	226 (37.9)	221 (37.1)	203 (34.0)	90 (16.4)	57 (10.7)	8 (1.8)	1 (0.2)
Infection	118 (19.8)	93 (15.6)	44 (7.4)	43 (7.2)	64 (11.7)	51 (9.5)	28 (6.3)	10 (2.2)
TLS	6 (1.0)	3 (0.5)	6 (1.0)	3 (0.5)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Thrombocytopenia	36 (6.1)	16 (2.7)	35 (5.9)	16 (2.7)	3 (0.5)	0 (0.0)	1 (0.2)	0 (0.0)
Acute thrombocytopenia	5 (0.8)	0 (0.0)	5 (0.8)	0 (0.0)	1 (0.2)	0 (0.0)	0 (0.0)	0 (0.0)
Haemorrhagic events	5 (0.8)	7 (1.2)	1 (0.2)	1 (0.2)	2 (0.4)	3 (0.6)	2 (0.5)	3 (0.7)
GI perforation	3 (0.5)	0 (0.0)	1 (0.2)	0 (0.0)	2 (0.4)	0 (0.0)	0	0
Cardiac events	22 (3.7)	17 (2.8)	11 (1.8)	7 (1.2)	9 (1.6)	9 (1.6)	2 (0.5)	1 (0.2)
Second malignancies*	30 (5.0)	17 (2.8)	0 (0.0)	0 (0.0)	19 (3.5)	15 (2.8)	12 (2.7)	2 (0.4)
Hepatitis B reactivation	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)

*No second malignancies were reported for induction period since these AEs are only captured 6 months after first study drug intake

Safety by chemotherapy subgroup

It should be noted that the GALLIUM study was not designed to compare induction chemotherapy regimens. The induction regimen was chosen on a per centre basis; therefore any differences between chemotherapies should be interpreted with caution.

In general, the baseline disease characteristics and demographics between the antibody arms within an individual chemotherapy regimen were comparable. The overall safety profile by chemotherapy subgroup is provided below.

Table 54: Summary of safety by chemotherapy group (safety population)

n, (%)	G-B (n=338)	G-CHOP (n=193)	G-CVP (n=61)	R-B (n=338)	R-CHOP (n=203)	R-CVP (n=56)
No. of patients with ≥1 AE	337 (99.7)	191 (99.0)	61 (100.0)	330 (97.6)	201 (99.0)	56 (100.0)
Total no. of events	5673	3357	1262	5236	3209	898
Total no. of deaths	26 (7.7)	7 (3.6)	2 (3.3)	32 (9.5)	9 (4.4)	5 (8.9)
No. of patients with ≥1:						
AE with fatal outcome	20 (5.9)	3 (1.6)	1 (1.6)	15 (4.4)	4 (2.0)	1 (1.8)
Grade 3–5 AE	231 (68.3)	170 (88.1)	40 (65.6)	224 (66.3)	151 (74.4)	30 (53.6)
SAE	171 (50.6)	74 (38.3)	26 (42.6)	155 (45.9)	64 (31.5)	19 (33.9)
SAE leading to treatment withdrawal	27 (8.0)	12 (6.2)	2 (3.3)	21 (6.2)	12 (5.9)	3 (5.4)
SAE leading to dose reduction	4 (1.2)	6 (3.1)	2 (3.3)	6 (1.8)	3 (1.5)	1 (1.8)
SAE leading to dose interruption	47 (13.9)	24 (12.4)	12 (19.7)	28 (8.3)	13 (6.4)	4 (7.1)
Related SAE	81 (24.0)	51 (26.4)	17 (27.9)	68 (20.1)	44 (21.7)	10 (17.9)
AE leading to treatment withdrawal	52 (15.4)	31 (16.1)	11 (18.0)	46 (13.6)	30 (14.8)	9 (16.1)
AE leading to dose reduction	43 (12.7)	51 (26.4)	13 (21.3)	46 (13.6)	38 (18.7)	11 (19.6)
AE leading to dose interruption	215 (63.6)	136 (70.5)	44 (72.1)	194 (57.4)	115 (56.7)	29 (51.8)
Related AE	317 (93.8)	183 (94.8)	61 (100.0)	304 (89.9)	192 (94.6)	51 (19.1)
Related AE leading to treatment withdrawal	38 (11.2)	23 (11.9)	11 (18.0)	30 (8.9)	27 (13.3)	8 (14.3)
Related AE leading to dose reduction	40 (11.8)	50 (25.9)	13 (21.3)	41 (12.1)	38 (18.7)	10 (17.9)
Related AE leading to dose interruption	186 (55.0)	121 (62.7)	42 (68.9)	167 (49.4)	105 (51.7)	24 (42.9)

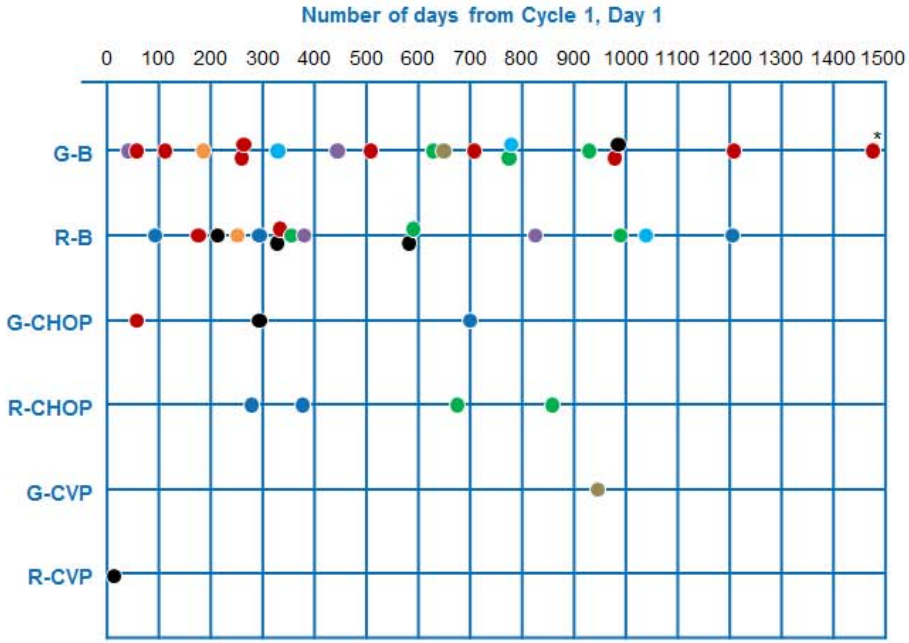
Treatment with bendamustine was associated with a higher incidence of Grade 3–5 infections and second malignancies during the maintenance and follow-up phases, while CHOP regimens were associated with higher rates of Grade 3–5 neutropenia during induction (Table 55). Furthermore, non-relapse fatal AEs were more common in bendamustine treated patients (G-benda 5.9% vs. R-benda 4.4%) than in those treated with CHOP (1.6% vs. 2.0%) or CVP (1.6% vs. 1.8%). The nature and timing of these events is shown below.

Table 55: Selected Grade 3–5 treatment-emergent AEs, listed by chemotherapy agent and treatment phase (safety population)

n, (%)	Induction						Maintenance						Follow-up					
	G			R			G			R			G			R		
Category	**B n=338**	**CHOP n=193**	**CVP n=61**	**B n=338**	**CHOP n=203**	**CVP n=56**	**B n=312**	**CHOP n=179**	**CVP n=57**	**B n=305**	**CHOP n=187**	**CVP n=43**	**B n=270**	**CHOP n=128**	**CVP n=44**	**B n=263**	**CHOP n=143**	**CVP n=45**
Neutropenia	73 (21.6)	124 (64.2)	24 (39.3)	87 (25.7)	103 (50.7)	13 (23.2)	49 (15.7)	36 (20.1)	5 (8.8)	29 (9.5)	26 (13.9)	2 (4.7)	6 (2.2)	2 (1.6)	0	1 (0.4)	0	0
Infections[†]	27 (8.0)	14 (7.3)	3 (4.9)	26 (7.7)	13 (6.4)	4 (7.1)	52 (16.7)	7 (3.9)	5 (8.8)	39 (12.8)	11 (5.9)	1 (2.3)	25 (9.3)	2 (1.6)	1 (2.3)	6 (2.3)	2 (1.4)	2 (4.4)
Second neoplasms[‡]	0	0	0	0	0	0	21 (6.7)	8 (4.5)	0	18 (5.9)	8 (4.3)	1 (2.3)	14 (5.2)	1 (0.8)	0	2 (0.8)	1 (0.7)	0

Figure 20: Incidence, nature and timing of non-relapse fatal AEs by chemotherapy agent and treatment arm in FL ITT Population

SOC ordered by decreasing frequency, n (%)	G-B, n=337	R-B, n=338	G-CHOP, n=191	R-CHOP, n=201	G-CVP, n=61	R-CVP, n=56
Infections and infestations	9 (2.7)	2 (0.6)	1 (0.5)			
Neoplasms benign, malignant, and unspecified	3 (0.9)	3 (0.9)	1 (0.5)	2 (1.0)		
General disorders and administration site conditions	1 (0.3)	3 (0.9)	1 (0.5)			1 (1.8)
Nervous system disorders		3 (0.9)		2 (1.0)		
Cardiac disorders	2 (0.6)	2 (0.6)				
Respiratory, thoracic, and mediastinal disorders	2 (0.6)	1 (0.3)				
Gastrointestinal disorders	1 (0.3)				1 (1.6)	
Metabolism and nutrition disorders	1 (0.3)	1 (0.3)				
Total	**19 (5.6%)**	**15 (4.4%)**	**3 (1.6%)**	**4 (2.0%)**	**1 (1.6%)**	**1 (1.8%)**



4.12.3 Overview of the safety of the technology in relation to the decision problem

The frequency and nature of AEs reported in GALLIUM was as expected for this study population (patients with FL) and for the treatment regimens being assessed. Overall, there were no new or unexpected safety findings with Gazyvaro in the first-line treatment of symptomatic patients with FL. The toxicity of G-chemo induction followed by extended treatment with Gazyvaro maintenance for 2 years was clinically manageable, as indicated by the high completion rate of dosing and limited number of dose delays and withdrawals due to AEs. However, non-relapse fatal AEs were more common in bendamustine-treated patients during all study phases, although absolute numbers were small

The incidence of AEs (all grades) was similar in the two treatment arms. The incidence of Grade 3–5 AEs was higher in the G-chemo+G arm compared with the R-chemo+R arm (74.6% vs. 67.8%). This was mainly due to a higher incidence of neutropenia AEs (43.9% vs 37.9%), febrile neutropenia (6.9% vs. 4.9%), IRRs (6.7% vs 3.7%), and thrombocytopenia (6.1% vs. 2.7%).

The incidence of SAEs was higher in the G-chemo+G arm compared with the R-chemo+R arm (46.1% vs. 39.9%). Adverse events leading to any dose modifications were also more frequent in the G-chemo+G (70.4% vs. 61.1%); the main drivers for this were neutropenia and IRRs.

The incidence of fatal AEs with G-chemo+G compared with R-chemo+R was similar in the two treatment arms (4.0% vs. 3.4%). In both treatment arms, most fatal AEs were infections or second malignancies. More deaths (for any reason, including progressive disease) occurred in the R-chemo+R arm (7.7%) than the G-chemo arm (5.9%).

Overall, the nature, frequency and severity of AEs of particular interest in patients with FL in this study were consistent with previous experience:

Infusion-related reactions

- The majority of IRRs were Grade 1 or 2 and there were no fatal IRRs.
- The overall incidence of IRRs was higher in the G-chemo+G arm (68.2% vs 58.5%), as was the incidence of Grade 3 and 4 IRRs, serious IRRs, and IRRs leading to withdrawal from treatment.
- The majority of IRRs occurred during Cycle 1, and IRRs decreased more dramatically from Cycle 2 onwards for all Grades, including Grade 3–4 AEs (Cycle 2:

15 AEs in the R-chemo+R arm vs. 8 AEs in the G-chemo+G arm), in the G-chemo arm+G and continued to decrease with subsequent cycles.

Neutropenia

- The incidence of neutropenia AEs was higher in the G-chemo+G arm (50.6%) than in the R-chemo+R arm (45.1%); this difference was driven mainly by neutropenia AEs in Cycle 1 (16.9% of patients in the R-chemo arm versus 23.0% in the G-chemo arm).
- Neutropenia AEs were most frequently observed during induction, and primarily during Cycles 1–6 (when chemotherapy was scheduled regardless of treatment arm).
- Concomitant G-CSF was administered to 45.8% of patients in the R-chemo+R arm and to 47.8% patients in the G-chemo+G arm, most commonly during Cycles 1-6.

Infection

- The overall incidence of infection (all treatment phases) was higher in the G-chemo+G arm (77.3%) than in the R-chemo+R arm (70.0%).
- The majority of infection AEs were Grade 1 or 2 in both treatment arms.
- The number of patients with Grade 3–5 infections was also higher in the G-chemo+G arm (20.0%) than in the R-chemo+R arm (15.6%), with two patients in the R-chemo+R arm and ten patients in the G-chemo+G arm experiencing fatal infections.
- In both treatment arms, a higher incidence of infections was observed during the maintenance phase compared to the induction phase. This may be explained by the longer duration of exposure to study treatment and/or the longer duration of observation of the maintenance (~2 years) compared to induction phase (~6 months).

Tumour lysis syndrome

- Tumour lysis syndrome was reported in 3/597 patients in the R-chemo+R arm and 6/595 patients in the G-chemo+G arm.
- All TLS events occurred during the first cycle of therapy.
- No fatal TLS was reported in either arm.
- No patient had to stop study treatment due to TLS, although 3 of the 6 patients in the G-chemo+G arm had study treatment temporarily interrupted due to TLS.

Thrombocytopenia

- The incidence of thrombocytopenia AEs was higher in the G-chemo+G arm (11.4%) than in the R-chemo+R arm (7.5%). The difference between arms was driven mainly by the AEs in Cycle 1.
- Of the patients who experienced thrombocytopenia events, one patient in the R-chemo+R arm and two patients in the G-chemo+G arm discontinued study treatment due to thrombocytopenia, and 13.3% of patients in the R-chemo+R arm and 30.9% patients in the G-chemo+G arm required treatment for this AE.
- Although a higher incidence of thrombocytopenia AEs was observed in the G-chemo+G arm, the incidence of hemorrhagic events was comparable (10.4% vs. 9.6%) between treatment arms, with very few Grade 3–5 AEs in either arm (1.2% vs. 0.8%).

Cardiac events

- The incidence of cardiac AEs was higher in the G-chemo+G arm (13.1%) compared with the R-chemo+R arm (9.7%).
- The majority of cardiac events were Grade 1 or 2.
- When excluding cardiac AEs reported as IRRs (such as palpitations, tachycardia, and bradycardia), the incidence of cardiac AEs was balanced between arms.
- The number of cardiac Grade 3–5 and serious AEs in patients without pre-existing cardiac conditions was low and balanced between arms.

Second malignancies

- The proportion of patients who experienced second malignancies starting 6 months or later after the first study drug intake was greater in the G-chemo+G arm (10.4%) compared with the with the R-chemo+R arm (7.0%).
- Non-melanoma skin cancers (basal cell carcinoma or squamous cell carcinoma) were the most frequently reported tumours (11 patients in the R-chemo+R arm and 16 patients in the G-chemo+G arm).
- Haematological malignancies were only reported in the G-chemo+G arm, but a variety of malignancies was reported (Hodgkin disease, acute myeloid leukaemia, acute lymphoblastic leukaemia) and no pattern was observed with regards to onset of the AE, latency or chemotherapy regimen.
- Solid tumours were also more frequently reported in the G-chemo arm but there was no clear difference between treatment arms in the incidence of any particular solid

tumour or group of tumours. No clear pattern was observed in the type of tumour, timing of onset of the AE, or latency in either treatment arm.

- There was no difference in fatal malignancies in the two arms (5 deaths in R-chemo+R arm and 6 deaths in the G-chemo+G arm).

Safety by treatment phase

The majority of AEs in both treatment arms occurred in the induction Phase in which the overall treatment intensity (antibody plus chemotherapy) is higher than during maintenance. During the induction period, the incidence of AEs (all grades, Grade 3–5, SAEs, and fatal) was comparable between the two treatment arms. Infusion-related reactions, neutropenia, thrombocytopenia, cardiac events, and TLS were more frequently reported during induction than in other phases in both treatment arms. During induction, cardiac events were mainly signs and symptoms of IRRs (e.g., tachycardia, bradycardia, palpitations). The incidence was higher in the G-chemo+G compared with the R-chemo+R arm for these AEs.

The overall safety in the maintenance phase was comparable between treatment arms. Infections were reported more frequently in maintenance than in other phases in both arms, and more frequently in the G-chemo+G arm.

During the follow-up phase, many fewer patients in both treatment arms experienced AEs than in other phases. Infections and neoplasms were the most frequently reported AEs.

Safety by chemotherapy regimen

GALLIUM was not designed to compare chemotherapy agents, nor were patients randomised to chemotherapy regimens; therefore, it is possible that there are differences in baseline characteristics between chemotherapy subgroups. Bearing these limitations in mind, bendamustine was associated with a higher rate of severe infections than CHOP or CVP during maintenance and follow-up. CHOP was associated with higher rates of early severe neutropenia, but this did not seem to translate into subsequent infection. Non-relapse fatal AEs were more common in bendamustine-treated patients during all study phases, although absolute numbers were small.

Safety profile summary

The toxicity of G-chemo+G was clinically manageable, as indicated by the high completion rate of dosing and the limited number of dose delays and withdrawals due to AEs, which is supported by the similar positive impact on QoL between the two treatment arms. Overall,

although the frequency of some AEs was higher in the G-chemo arm, no new or unexpected safety signals were detected with G-chemo+G in patients with FL.

4.13 Interpretation of clinical effectiveness and safety evidence

Briefly conclude the clinical effectiveness and safety of the technology against the comparators specified in the final scope issued by NICE, including any subgroups. If relevant, include a statement on whether this technology meets the end-of-life criteria. Complete the table below and cross reference to where this information is found in the company submission.

As discussed in Section 3.3, there remains an unmet need for some patients with FL, i.e. those who suffer early disease progression (approximately one fifth of FL patients receiving immunochemotherapy still suffer a PD event within two years); experience transformation following diagnosis or treatment initiation (associated with a significantly poorer OS than patients who transform later [22% vs. 66%]); and those in particular subgroups (e.g. high-risk FLIPI/FLIPI2 groups). Novel and effective therapies that target these high-risk patient subpopulations are needed.

Evidence for the efficacy and safety of Gazyvaro in patients with previously-untreated advanced FL is available from the Phase III open-label GALLIUM (BO21223) study. GALLIUM compared Gazyvaro in combination with chemotherapy followed by Gazyvaro monotherapy as maintenance (G-chemo+G) with MabThera in combination with chemotherapy followed by MabThera monotherapy as maintenance (R-chemo+R), which is regarded as the standard of care for first-line treatment of advanced, symptomatic FL.

Summary of clinical efficacy

At the pre-planned GALLIUM interim analysis, G-chemo+G demonstrated a clinically meaningful and statistically significant reduction of 34% of the risk of investigator-assessed PFS compared with R-chemo+R (stratified HR 0.66 [95% CI: 0.51, 0.85; p=0.0012). On the basis of KM estimates, 80.9% (95% CI, 77.4%, 84.0%) and 73.3% (95% CI: 68.8, 77.2) of patients in the R-chemo+R arm were progression-free at two and three years, respectively, compared with 87.7% (95% CI, 84.6%, 90.1%) and 80.0% (95% CI: 75.9, 83.6) of patients in the G-chemo+G arm. The results of the IRC assessment of PFS were consistent with the investigator-assessed PFS results (stratified HR 0.71 [95% CI 0.54; 0.93]; p=0.0138), while other secondary time-to-event endpoints (OS, EFS, DFS, DoR, and NALT) were supportive of the PFS outcomes.

Reflecting the indolent nature of FL disease, and after a median follow-up of approximately 34.5 months, median PFS was not expected to be reached at interim analysis. Based on the PRIMA study, where 59.2% of previously untreated FL patients on R maintenance were progression-free at 6 years after 73 months' median follow-up (Seymour JF et al., 2013) and assuming a conservative median PFS of six years for R-chemo+R, the observed HR of 0.66 in GALLIUM would translate to a 1.5x longer median PFS for G-chemo+G than R-chemo+R, and to an estimated three year improvement in the G-chemo+G arm. Longer follow-up data will confirm if these benefits are achieved.

In indolent cancers such as FL, punctuated by a series of remissions and relapses, patients may survive for many years despite PD. As death is a less common outcome than PD, improved PFS would not be expected to translate into a significant OS benefit after two to three years of follow-up. OS was a secondary endpoint, as such, GALLIUM was not powered to detect a difference in OS between the two antibody treatment groups. Nevertheless, after a median follow-up of 34.5 months, with <20% of patients followed for OS for more than four years, the HR for OS was 0.75 (95% CI, 0.49–1.17, p=0.21). Based on KM estimates, the estimated probability of being alive at three years was 92.1% (95% CI, 89.5–94.1) in the R-chemo+R arm and 94.0% (95% CI, 91.6–95.7) in the G-chemo+G arm. On visual inspection, the KM plot for OS showed a separation of the curves favouring the G-chemo+G arm.

Several pre-specified subgroup analyses showed that the investigator-assessed PFS benefit with G-chemo+G was consistent across all patient subgroups. With the exception of FL FLIPI low risk (HR 1.17 [95% CI: 0.63, 2.19]; based on 253 patients), the observed hazard ratios were below 1.00 and ranged from 0.40–0.86 for subgroups including at least 10% of patients. GALLIUM was not designed to compare the three different chemotherapy regimens used in the study (CHOP, CVP or bendamustine). As the allocation of chemotherapy was not randomised at the patient level, there may be confounding differences in baseline patient characteristics between the chemotherapy subgroups. Pre-planned subgroup analyses of investigator-assessed PFS HRs showed that all G-containing chemotherapy regimens had a consistent benefit over R-chemo regimens in FL patients (CHOP, 0.77 [95% CI, 0.50–1.20]; CVP, 0.63 [95% CI, 0.32–1.21]; bendamustine, 0.61 [95% CI, 0.43–0.86]).

Summary of safety

The current standard of care for previously-untreated symptomatic FL, MabThera plus chemotherapy followed by MabThera maintenance, is associated with clinically manageable AEs. In GALLIUM, the toxicity of G-chemo+G was clinically manageable, as indicated by the

high completion rate of dosing and the limited number of dose delays and withdrawals due to AEs, which is supported by the similar impact on QoL between the two treatment arms. Furthermore, the high rate of treatment completion, and the limited number of chemotherapy dose reductions indicate that G-chemo+G was generally well tolerated.

The nature of AEs observed were consistent with the known profiles of the study treatments, with a similar incidence of all grade AEs in the two arms; 98.0% of patients in the R-chemo+R arm vs. 99.5% of patients in the G-chemo+G arm. While patients in the G-chemo+G arm had a numerically higher frequency of grade 3 to 5 AEs and SAEs than patients in the R-chemo+R arm, the rate of fatal (grade 5) AEs was comparable between the treatment arms. Overall, although the frequency of some AEs was higher in the G-chemo+G arm, no new or unexpected safety signals were detected.

Clinically relevant IRRs of grade 3 or higher occurred in 6.7% of G-chemo+G patients, which is similar to the values reported in GADOLIN (MabThera relapsed/refractory FL) (Sehn et al., 2016, Sehn et al., 2015), and less frequent than CLL patients with comorbidities (Goede et al., 2014).

Bendamustine was associated with higher rates of severe infections than CHOP or CVP during maintenance and follow up in both treatment arms. Non-relapse fatal AEs were also more common in bendamustine-treated patients during all study phases, although absolute numbers were small. The MRD data from GALLIUM provide evidence that less intensive chemotherapy regimens combined with Gazyvaro still demonstrate greater efficacy than when given with MabThera and maintain the overall beneficial effect of Gazyvaro (Pott C et al., 2016).

Strengths and limitations of clinical evidence

The study population in GALLIUM is largely reflective of the advanced FL population in the UK. More patients were recruited from the UK than any other country (293 patients from 29 centres), indicating that the results of GALLIUM will reflect UK practice. Furthermore, feedback from clinical experts confirms that the baseline characteristics of FL patients enrolled into GALLIUM are reflective of the population seen in UK clinical practice. It has been noted however that the time from diagnosis to treatment is shorter compared with clinical practice; a higher proportion of patients receiving treatment soon after diagnosis could be indicative of a more aggressive cohort. Furthermore, the chemotherapy regimens included in GALLIUM reflect current UK clinical practice, as demonstrated by the SACT dataset for UK chemotherapy regimens used in 2014, presented in Table 13.

Gazyvaro is compared against a relevant active comparator in GALLIUM as R-chemo followed by MabThera maintenance therapy is regarded as the standard of care for the first-line treatment of patients with advanced FL. Furthermore, GALLIUM was designed to capture endpoints which are relevant to UK clinical practice and that address the unmet medical need for this patient population.

PFS was assessed both by the investigator (primary endpoint) and IRC. Concordance between the investigator and IRC assessment of PD was analysed in terms of the type of event (i.e. PD event or death). Agreement on the type of event was high (92.1%) overall and balanced between arms (91.0% in the R-chemo+R arm vs. 93.2% in the G-chemo+G arm). In particular, for the R-chemo+R arm, the proportion of patients who were assessed as progression-free by the investigator and IRC was 73.9% and 79.2%, respectively, and patients reported to have disease progression was 15.1% and 17.6%, respectively. Moreover, for the G-chemo+G arm, the proportions of patients assessed to be progression-free by the investigator (80.9%) and the IRC (84.5%) and patients reporting disease progression as assessed by the investigator (9.0%) and IRC (11.4%) were similar.

Concordance/discordance in the timing of PFS event as determined by the investigators and IRC assessments was also analysed. Overall, IRC-assessed and investigator-assessed timing of PD were largely in agreement (within 30 days of each other). In cases where the IRC- and investigator assessed date of PD differed, the difference in timing was similar in the two treatment arms, suggesting that there was no systematic bias attributable to investigators' knowledge of individual patient's treatment allocation.

As GALLIUM was not designed to evaluate treatment benefits separately for the induction and maintenance phases, it is not possible to determine whether the higher PFS rate with G-chemo+G resulted from any one particular stage of the study. However, it is very unlikely that Gazyvaro would provide the PFS benefit observed in GALLIUM if it was used as maintenance only. For instance, in the EORTC-20981 study of patients with relapsed/refractory FL, median PFS with MabThera as induction and maintenance was 4.4 years, compared with only 3.1 years for those who received MabThera maintenance only (van Oers et al., 2010). In addition, in GALLIUM, MRD-negativity and CR rates with PET at the EOI were significantly higher in the G-chemo arm, which suggests that Gazyvaro may induce deeper responses than MabThera during induction. MRD can offer additional information regarding antibody efficacy, particularly with respect to changes in tumour burden. In an exploratory analysis of MRD, a significantly greater proportion of patients in the G-chemo arm achieved MRD-negative status in PB at mid-induction (94.3% vs. 88.9%;

p=0.0132) and in PB and/or bone marrow at the EOI (92.0% vs. 84.9%; p=0.0041) compared with patients in the R-chemo arm. These findings suggest that G-chemo based induction may induce more rapid and more effective tumour-cell clearance than R-chemo based treatment.

In conclusion, GALLIUM demonstrates that replacing MabThera with Gazyvaro in the immunochemotherapy induction and monotherapy maintenance setting for previously untreated FL patients produces a meaningful improvement in PFS. Although the frequency of some AEs was higher with Gazyvaro, no new safety signals were detected and the benefit/risk ratio remains positive. G-chemo+G therefore represents a significant improvement in therapy for this patient population.

End-of-life criteria

This technology does not meet the end-of-life criteria because patients with FL are expected to have life expectancy beyond 24 months (Table 56).

Table 56: End-of-life criteria

Criterion	Data available
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	No, the median OS in patients with untreated patients with advanced FL is greater than 24 months
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	No, at the time of clinical cut-off, less than 20% of patients had been followed for survival for more than 4 years, hence the data can be considered still immature at this time (stratified HR for overall survival: 0.75 [95% CI:0.49, 1.17], stratified log-rank p=0.21).
The treatment is licensed or otherwise indicated for small patient populations	In 2015, 2,142 new cases of FL were registered in England (Office for National Statisitics, 2017). Estimated number of patients that will be treated with first line immunochemotherapy induction is highlighted in Section 6.2 (1,152 patients).

4.14 Ongoing studies

The GALLIUM study is ongoing. Further analysis from an updated data cut (clinical cut-off 16th September 2016) that formed the basis of the economic analysis will be available within the next 12 months, as well as a 90-day safety update for the FDA. There are plans to present follow up analyses from the available data cuts of GALLIUM at international conferences in 2017, including:

- PET analysis (ICML 2017)
- Analysis by chemotherapy regimen (ICML 2017)
- Health-related QoL data (EHA 2017)

There are no further studies ongoing investigating Gazyvaro in the apprased indication.

5. Cost effectiveness

5.1 Published cost-effectiveness studies

5.1.1 Identification of studies

Describe the strategies used to retrieve cost-effectiveness studies relevant to decision-making in England from published NICE technology appraisals, the published literature and from unpublished data held by the company. Justify the methods used with reference to the decision problem and the NICE reference case. Provide sufficient detail to enable the methods to be reproduced, and the rationale for any inclusion and exclusion criteria used. Provide the search strategy used in an appendix

Nomenclature used for GALLIUM in the cost-effectiveness section:

● Gazyvaro (obinutuzumab; G) or MabThera (rituximab; R) in combination with chemotherapy as induction therapy, followed by G or R monotherapy as maintenance is abbreviated as G-chemo+G and R-chemo+R, respectively. G-chemo+G represents the regimen as per the anticipated Marketing Authorisation

Search strategy development

The aim of the strategy was to identify studies of economic evaluations of treatments in the first-line FL setting that could inform the de novo economic analysis.

The search strategy was developed using a combination of free text, MEDLINE MeSH and EMBASE terms, as appropriate for the databases included. Briefly, the search terms in the strategy included:

- Disease state terms for iNHL
- Line of treatment (i.e. previously untreated)
- Cost, resource use, HRQoL or health state utility (HSUV) terms.

Further details are shown in Appendix 5.

Data sources

Electronic databases searched included MEDLINE, MEDLINE In-Process, EMBASE, the Cochrane Library and NHS EED. Hand searches were conducted in conference abstracts including the following organisations:

- American Society of Hematology (ASH)
- American Society of Clinical Oncology (ASCO)
- European Society of Medical Oncology (ESMO)
- European Haematology Association (EHA)
- International Conference on Malignant Lymphoma (ICML)
- International Network of Agencies for Health Technology Assessment (INAHTA)
- National Institute for Health and Care Excellence (NICE)
- National Institute for Health Research HTA
- Canada: Canadian Agency for Drugs and Technologies in Health (CADTH)

Search implementation date and span

Initial searches were conducted on 14th June 2016 with hand searches in May 2016. Studies from 1998 onwards were considered as this coincided with the market approval date of MabThera. Hand searches covered the past three years from the search date. The searches were then updated on 7th March 2017 with hand searches also in March 2017.

Study selection process

Screening

All citations were screened initially by an analyst and then screened in a blinded manner by a second analyst. Any discrepancies were resolved by discussion and a third reviewer was consulted to resolve disagreements. Studies were screened for exclusion from the review of economic evaluations or utility studies (see section 5.3) using an adaptation of the PICOS framework in Table 57. Briefly, studies with any relevant economic outcomes were considered, and intervention and comparators were not restricted.

Table 57: Criteria at screening stage for full text review of economic evaluations and utility studies

PICOS	Definition
Population	People in the UK with iNHL who were previously untreated. All subtypes, except skin lymphomas
Intervention & Comparators	• Not restricted • Any intervention (transplantation included, therapies aimed at specific comorbidities excluded)
Outcomes	The outcome measures to be considered for the economic literature review are: • Costs • Resource use • Quality of life • Utility

Study types	• Health economic evaluations for economic endpoints

Inclusion and exclusion criteria

The inclusion criteria were defined using an adaptation of the PICOS framework in Table 58 below and applied at the full text review stage.

Table 58: Inclusion and exclusion criteria for economic evaluations

PICOS	Inclusion criteria	Exclusion criteria
Population	• People in the UK with iNHL who were previously untreated. • All subtypes, except skin lymphomas	• Disease area not iNHL • Relapsed or refractory setting • Setting not UK
Intervention & Comparators	• Intervention and comparator not restricted	
Outcomes	The outcome measures to be considered for the economic literature review are: • Costs • Resource use • Quality of life • Utility	
Study types	Health economic evaluations	Other study types: • Secondary publications • Review articles, systematic literature reviews, or meta-analyses • Editorials, notes or letters to the editor • Studies containing no primary data

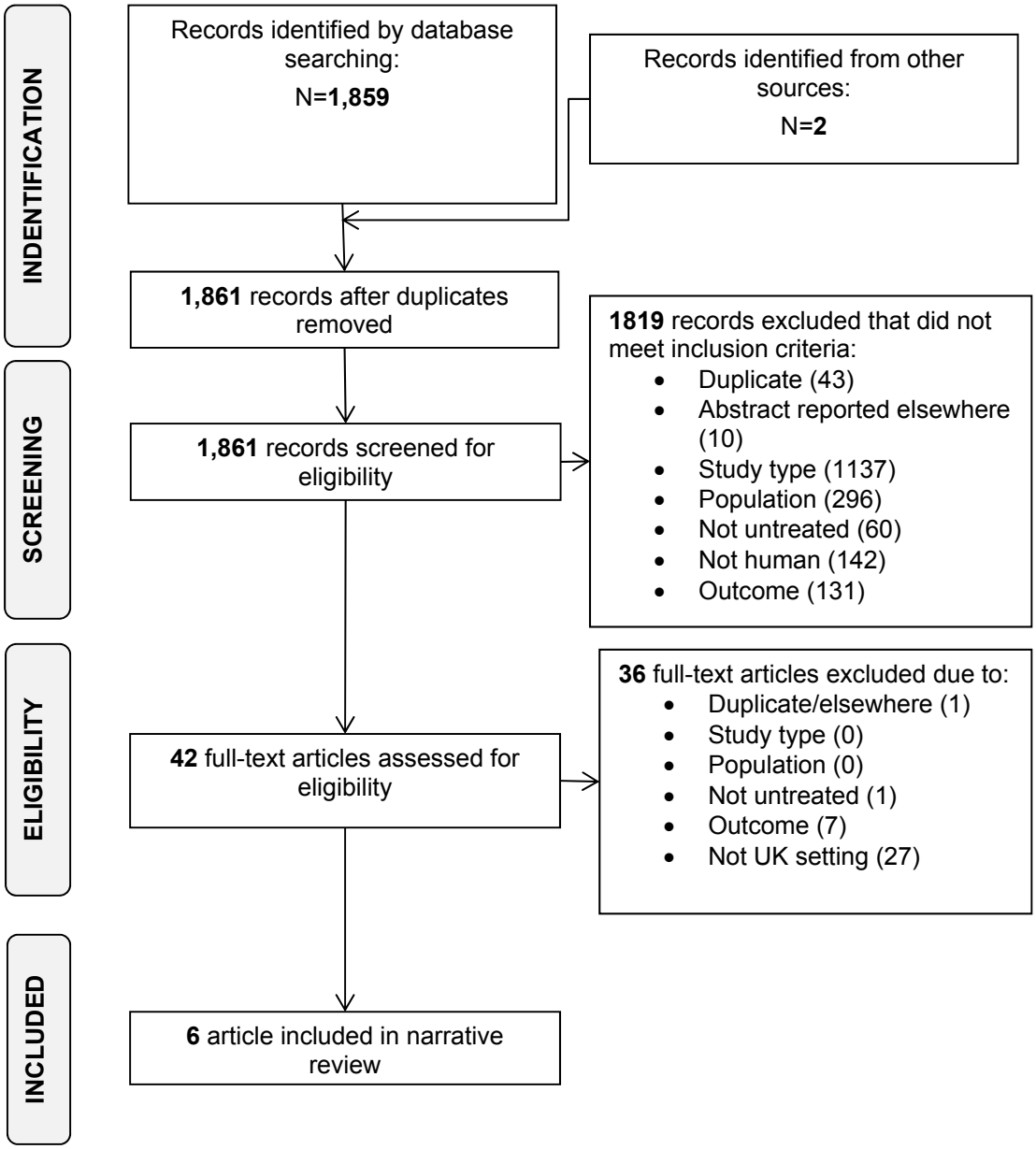
Results

The PRISMA flow-diagram outlining the study selection process is presented in Figure 21 with numbers combining the original and updated searches.

The search strategy identified a total of 1,861 records from the electronic databases and from supplementary searching after removal of duplicates. 1,819 studies were excluded at screening due to duplicates (43), abstracts being reported elsewhere (10), study type (1,137), population (296), not untreated (60), not human (142) or outcome (131).

Of the 42 studies reviewed at the full text stage, 6 UK studies were included in the narrative review. 27 were non-UK studies, 7 were excluded due to outcome reported, 1 was not relevant for previously-untreated patients and one was a duplicate.

Figure 21: PRISMA diagram for cost-effectiveness studies



5.1.2 Description of studies

Provide a brief overview of each cost-effectiveness study only if it is relevant to decision-making in England. Describe the aims, methods and results for each study. Each study's results should be interpreted with reference to a critical appraisal of its methodology. When studies have been identified and not included, justification for this should be provided. If more than 1 study is identified, please present the information in a table as suggested below

None of the studies identified in the systematic literature review addressed the decision problem as no study investigated the cost effectiveness of G-chemo+G as an intervention in

previously-untreated patients with FL; this also included studies from other countries that were excluded at full text review. The UK studies included in Table 59 below were cost-effectiveness studies in the first-line settig of FL of rituximab in combination with various chemotherapy regimens as induction (Dundar et al., 2009, Papaioannou et al., 2012, Ray JA et al., 2010), as maintenance (Greenhalgh et al., 2013), as induction followed by maintenance (Dewilde et al., 2014), or a study deriving costs and outcomes from a UK observational cohort via a simulation model (Wang H et al., 2016).

Due to the indolent nature of the disease, long term progression-free or overall survival data over the patient's lifetime is not available from a single trial. Therefore, studies either used Markov models ((Dundar et al., 2009, Greenhalgh et al., 2013, Ray JA et al., 2010) or microsimulation approaches (Dewilde et al., 2014, Papaioannou et al., 2012) and include outcomes and costs from further treatment lines to estimate overall costs and outcomes in intervention and comparator arms.

Table 59: Summary of UK cost effectiveness studies in previously untreated FL

Study	Year	Patient population	Summary of model	Intervention - comparator	Costs (intervention, comparator)	QALYs (intervention, comparator)	ICER (per QALY gained)
(Dundar et al., 2009)	2009	NHL, Stage III/IV	Markov (PFS, PD & Death)	R-CVP vs CVP alone	N/R	N/R	<30,000
(Ray JA et al., 2010)	2010	FL, Advanced disease	Markov (PFS, PD & Death)	R-MCP vs MCP R-CVP vs CVP R-CHOP vs CHOP R-CHVP vs CHVP	£29,725vs£20,900 £28,582 vs £20,708 £29,794 vs £20,922 £33,513 vs £29,621	6.75 vs 5.56 5.39 vs 4.75 6.34 vs 5.50 5.97 vs 5.51	7,454 8,614 10,676 4,683
(Papaioannou et al., 2012)	2012	Symptomatic Stage III/IV	Patient level simulation (PFS1, PFS2, PD & Death)	R-MCP vs MCP R-CVP vs CVP R-CHOP vs CHOP	41,370 vs 36,103 38,183 vs 30,793 40,708 vs 34,983	7.36 vs 6.79 6.95 vs 5.99 7.37 vs 6.84	9316 7,720 10,834
(Greenhalgh et al., 2013)	2013	FL, Advanced disease	Markov (PFS1, PFS2, PD & Death)	R vs Observation (maintenance)	£70,666 vs £52,823	7.87 vs 6.83	17,136
(Dewilde et al., 2014)	2014	iNHL	Patient level simulation (PFS1, PFS2, PD & Death)	B-R+R vs R-CHOP+R or R-CVP+R	£63,453 vs £59,627 or £58,532	7.19 vs 6.46 or 6.58	B-R vs R-CHOP: 5,249 B-RIT vs R-CVP:£8, 092/QALY
(Wang H et al., 2016)	2016	FL	Patient level simulation	Chemotherapy +/- rituximab, radiotherapy, watch and wait.		N/R	N/R

5.1.3 Provide a complete quality assessment for each relevant cost-effectiveness study identified. Use an appropriate and validated instrument, such as those of Drummond and Jefferson (1996)[2] or Philips (2004)[3]. Please provide these assessments in an appendix.

See appendix 5.

5.2 De novo analysis

5.2.1 Patient population

State which patient groups are included in the economic evaluation and how they reflect the population defined in the scope and decision problem for the NICE technology appraisal, Marketing Authorisation/CE marking, and the population from the trials. If there are differences, please provide the rationale. Explain the implications of this for the relevance of the evidence base to the decision problem. For example, indicate if the population in the economic model is different from that described in the (draft) summary of product characteristics (SmPC) or information for use (IFU) and included in the trials

The patient population of the de novo economic analysis is based on the FL trial population in GALLIUM which equates to the expected license indication and place in clinical practice (Section 3).

These are patients with advanced FL who require treatment. The population is in line with the decision problem set out in Section 1.2. The patient disposition of the GALLIUM study is discussed in detail in Section 4.5 and the key demographic variables for the model are based on the GALLIUM trial FL population as summarised in Table 60 below.

Table 60: Model demographic variables based on GALLIUM

Variable	Value
Average age of cohort (years)	57.9
Body weight (kg)	75.7
Height (cm)	168.3
Calculated Body Surface Area (m^2)	1.86

5.2.2 Model structure

Describe the model structure and provide a diagram of the model submitted, including the following:

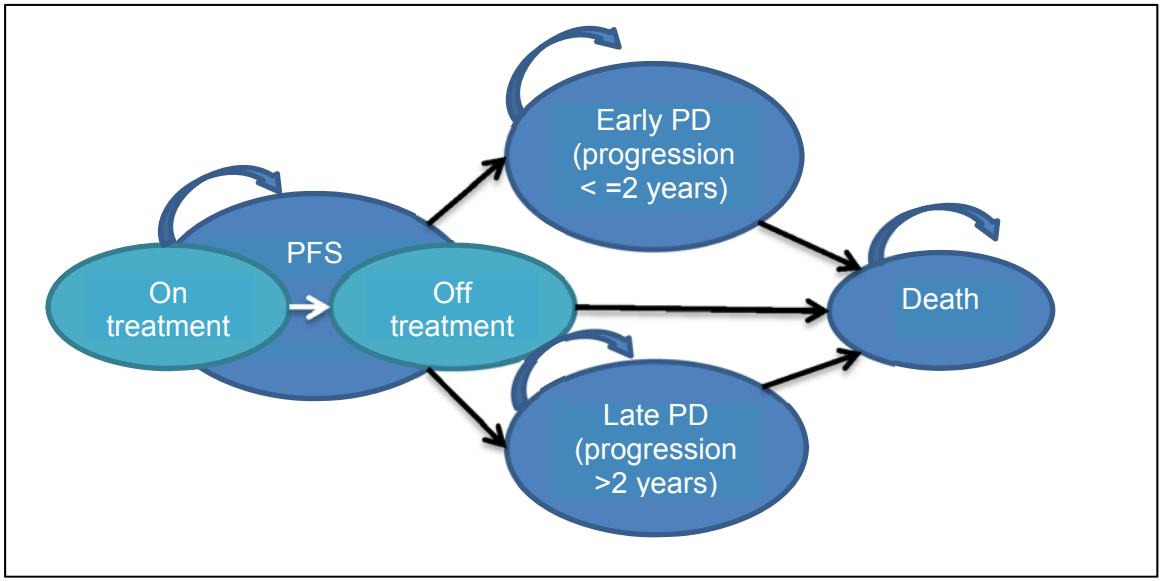
- **Type of de novo analysis (for example, decision tree, Markov model, discrete event simulation model).**
- **Justification of the chosen structure in line with the clinical pathway of care described in section 3.3.**
- **How the model structure and its health states capture the disease or condition for patients identified in section 3.3.**
- **Where appropriate, state the cycle length and whether a half-cycle correction has been applied.**

A four-state Markov model was developed (Figure 22) with a progression-free state (PFS) (on/off treatment) and two progressed disease (PD) states, early PD and late PD (with subsequent treatments) and death.

As outlined in Section 5.3.1 and Section 3 the time to progression after initial treatment is highly predictive for post-progression mortality and overall survival. In particular patients progressing early, i.e., within two years of initial treatment, have significantly worse mortality than patients who did not progress within two years (Casulo et al., 2015b).

The two PD states may include multiple lines of treatment post-progression. Outcomes and costs of these later treatment lines are accounted for by average cost and mortality and do not require specific treatment sequence assumptions. Outcomes for patients experiencing early progression in the model were based on GALLIUM data as the follow-up period mainly includes the early progression events (up to two years after initial treatment). Outcomes for patients experiencing late progression were based on data from the PRIMA study. PRIMA is the main Phase III, randomised controlled trial of rituximab maintenance in patients with high tumour burden FL responding to R-chemo induction (Salles et al., 2011, Seymour JF et al., 2013). Patient level data was available to construct an R-chemo+R arm with long term follow up (up to 9.75 years) by combining data from the induction phase with R-chemo (pre-randomisation in PRIMA) with the data from the R-maintenance arm (see Appendix 6 for details).

Figure 22: De novo model structure



PD, progressed disease; PFS, progression-free survival

As briefly described in Section 5.2.1, previous models for the cost effectiveness of rituximab in combination with chemotherapy induction were Markov models with a PFS state, a progressed disease state and death (Ray JA et al., 2010). A separate state for second-line treatment and remission (PFS2) was also considered (Greenhalgh et al., 2013). Alternatively, microsimulation approaches allowed (Dewilde et al., 2014, Papaioannou et al., 2012) for more complex treatment sequences, e.g. choice of second-line treatment depending on first-line treatment and response. Similar to the de novo model described here, these models rely on various data sources to estimate transitions and outcomes on different lines of treatment. However, these approaches did not explicitly account for the striking correlation between time to progression and overall survival outcomes discussed in Section 5.3.1. The current de novo model was therefore chosen to incorporate these findings and present a model structure that incorporates the outcomes of interest as well as their correlation in a straight forward way for decision making.

PFS (on treatment and off treatment)

Initially all patients begin in the PFS health state on treatment (G-chemo+G or R-chemo+R) and are assigned a PFS 'on-treatment' utility value and treatment costs while on therapy. Patients are treated in a similar manner in both arms, the only difference being the anti-CD20 therapy (Gazyvaro or MabThera) administered with chemotherapy induction and as maintenance therapy for responders. During active treatment, patients receive additional supportive care/monitoring as described in section 5.5.

Time on treatment was determined using patient level data from GALLIUM for both arms. The model uses the observed Kaplan-Meier time-to-treatment-discontinuation (TTTD) curves for individual R-chemo+R and G-chemo+G strata to estimate the proportion on treatment in each cycle of the model. Extrapolation of TTTD was not required as the data was mature, i.e. patients had completed treatment in both arms.

When patients complete or discontinue treatment in the PFS state, they are considered off treatment and assigned an 'off treatment' PFS utility value and costs for ongoing monitoring in supportive care as described in section 5.5. Separate values for utilities and costs were used for the induction and maintenance phase. Patients can either remain in PFS (on- or off-treatment) or exit the state due to disease progression or death.

PD states

On progression during or after first line treatment patients move to the progressed disease (PD) health states at any time. Patients progressing within two years of treatment have significantly worse outcomes compared to patients progressing later (Casulo et al., 2015b, Maurer et al., 2016) (see Section 5.3.1). To be able to apply different outcomes and costs to the cohorts of patients who experience an early or a late progression, two progressed disease states were introduced (early PD and late PD). Patients enter the respective PD states according to the time of progression in the model; patients progressing within two years from the beginning of the initial treatment enter the early PD state and patients progressing after two years enter the late PD state, respectively. Once patients enter any of the two PD states, patients cannot transition back to PFS. In addition, patients entering the early PD state stay in this state until death and cannot transition to the late PD state. Patients in the late PD state remain in this state until death and cannot transition to early PD state. Transitioning between the two PD states (early PD and late PD) is not possible given that the two states are both mutually exclusive and independent.

Death state

Patients move into the death state at any time from either the PFS or the PD health states. The death state is an absorbing state; the proportion of patients in this state is calculated by the sum of deaths in the PFS and PD states. The cumulative deaths from PFS, early and late PD states are used to calculate overall survival in the model.

5.2.3 Features of the de novo analysis

Complete the table below presenting the features of the de novo analysis. Compare and justify your chosen values with the methods specified by NICE in the reference case

The features of the de novo analysis are summarised below.

Table 61: Features of the de novo analysis

Factor	Chosen values	Justification
Time horizon	Lifetime (equating to a maximum of 40 years)	NICE reference case. Approximately 1% of patients were alive in the R-chemo+R arm at 40 years
Were health effects measured in QALYs; if not, what was used?	QALYs	NICE reference case
Discount of 3.5% for utilities and costs	3.5%	NICE reference case
Perspective (NHS/PSS)	NHS/PSS	NICE reference case
Cycle length	1 month	Appropriate to cover treatment cycles and outcomes over the time horizon for indolent disease
Half-cycle correction	Yes, applied to all Markov traces	NICE reference case

PSS, personal social services; QALYs, quality-adjusted life years

Intervention technology and comparators

5.2.4 If the intervention and comparator(s) are not implemented in the model as per their Marketing Authorisations/CE marking, describe how and why there are differences. Make it clear whether the intervention and comparator(s) included in the model reflect the decision problem. If not, briefly describe how and why, cross-referencing to the decision problem section in your submission

The intervention and comparator are in line with the decision problem set out in section 1.4. G-chemo+G was implemented as per the anticipated Marketing Authorisation in the intervention arm as set out in the clinical Section 4.3.1. In the comparator arm, R-chemo+R was implemented as per Marketing Authorisation and current clinical practice. The relevance of R-chemo+R and the different chemotherapy options as the relevant comparator in clinical practice is discussed in detail in Section 3.

More details on the implementation of the technologies within the models can be found in Section 5.5.

5.2.5 If a treatment continuation rule has been assumed for the intervention and comparator(s), provide the rationale for the continuation rule and where it is referenced (for example, [draft] SmPC, European public assessment report, comparator use, clinical practice, or clinical trial protocols). Please note that this refers to clinical continuation rules and not patient access schemes. If a treatment continuation rule is included in the model that is not stated in the (draft) SmPC or information for use (IFU), this should be presented as a separate scenario by considering it as an additional treatment strategy alongside the base-case interventions and comparators. Consideration should be given to the following:

- **the costs and health consequences of implementing the continuation rule (for example, any additional monitoring required)**
- **the robustness and plausibility of the end point on which the rule is based**
- **whether the 'response' criteria defined in the rule can be reasonably achieved**
- **the appropriateness and robustness of the time at which response is measured**
- **whether the rule can be incorporated into routine clinical practice**
- **whether the rule is likely to predict those people for whom the technology is particularly cost effective**
- **Issues about withdrawal of treatment for people whose disease does not respond and other equity considerations.**

Treatment continuation rules have not been applied in the economic model. Time to treatment discontinuation is based on the actual observation from the GALLIUM study for both arms. Specifically, as per license indication, only patients responding to induction received maintenance. Maintenance was only offered until progression or for a maximum of two years.

5.3 Clinical parameters and variables

5.3.1 Describe how the clinical data were incorporated into the model, also commenting on the following factors:

- **Whether intermediate outcome measures were linked to final outcomes (for example, if a change in a surrogate outcome was linked to a final clinical outcome). If so, explain how the relationship was estimated, what sources of evidence were used, and what other evidence there is to support it.**
- **Whether costs and clinical outcomes are extrapolated beyond the trial follow-up period(s). If so, explain and justify the assumptions that underpin this**

extrapolation, particularly the assumption that was used about the longer-term difference in effectiveness between the intervention and its comparator. For the extrapolation of clinical outcomes, present graphs of any curve fittings to patient-level data or Kaplan–Meier plots and the methods and results of any internal and external validation exercises. The NICE Decision Support Unit[4] has published technical support document 14, which provides additional information on the implementation of methods and reporting standards for extrapolation with patient level data.

Clinical parameters for the model were derived from the GALLIUM trial data for PFS and post-progression survival (PPS) for early progression (PPS in early PD). External data was used to populate the PPS for late progression using long term data from PRIMA. The latest available data cut of GALLIUM with a clinical cut-off date of 16th September 2016 was used.

The investigator (INV) assessed PFS data (PFS-INV) was used, corresponding to the primary endpoint (see Sections 4.4. and 4.7). The use of independent review committee (IRC) assessed PFS (PFS-IRC) was investigated in a sensitivity analysis. The extrapolation beyond the observed period in the GALLIUM trial was based on parametric functions as described below.

To derive PPS for patients progressing late, data sources with longer follow up than GALLIUM were required to obtain sufficient death events for this group. Data from the PRIMA study was used in the base case to estimate the mortality post progression for late PD as this data was based on a cohort receiving R maintenance after response to R-chemo induction treatment where patient level data with up to 9.75 years of follow up was available. However, as described in Appendix 6 a R-chemo+R cohort had to be constructed from patient level data for patients randomised to maintenance (PRIMA patients), that allowed estimates for PFS and PPS from the start of R-chemo induction therapy (as in GALLIUM).

The transitions used in the model and the data sources are summarised in Table 62 and are discussed in more detail below.

Table 62: Summary of the health state transitions used in the model

Transition	Transition probability	Source in submission
PFS to early PD and late PD	Time dependent calculated from the probability of remaining in PFS and probability of death in PFS. Probability of remaining in PFS modelled with parametric model (base case Weibull) and proportional hazards.	Table 65
PFS to death	Mortality rates based on trial mortality in GALLIUM and general population background mortality.	Table 66
Early PD to death	Post-progression mortality for early progression based on GALLIUM mortality.	Table 67
Late PD to death	Post-progression mortality for late progression based on PRIMA mortality.	Table 67

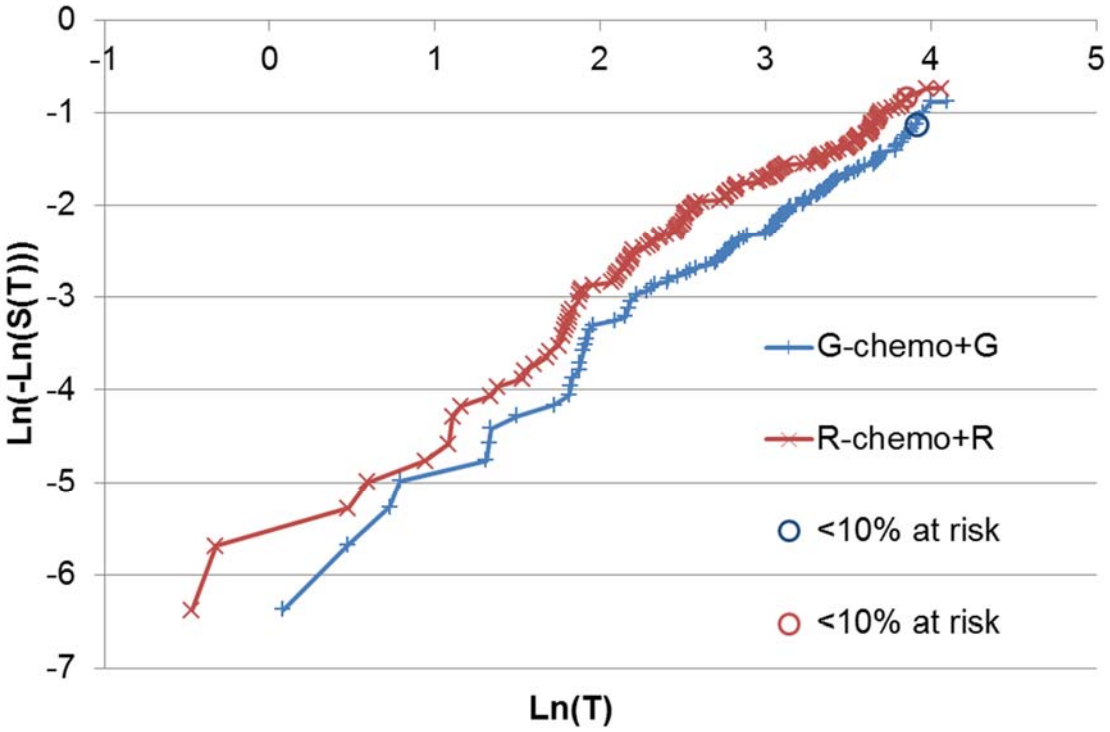
Probability of remaining in PFS

In the model, the probability of remaining in PFS is determined from a parametric function fitted to patient level PFS-INV data from GALLIUM (see also section 4.7). In electing the appropriate function the NICE Decision Support Unit (DSU) guidance was followed (Latimer, 2013). Several commonly used parametric distributions were fitted to individual patient level PFS data and investigated for suitability to extrapolate beyond the observation period based on visual inspection, goodness of fit and external validity. Functions investigated were Exponential, Weibull, Log-logistic, Log-normal, Gamma and Gompertz. Data was relatively immature with 26.8% and 20.0% (24.0% and 16.8% in the primary analysis) of patients having progressed or died in the R-Chemo+R and G-Chemo+G arm, respectively.

Proportional hazards assumption of PFS parametric functions

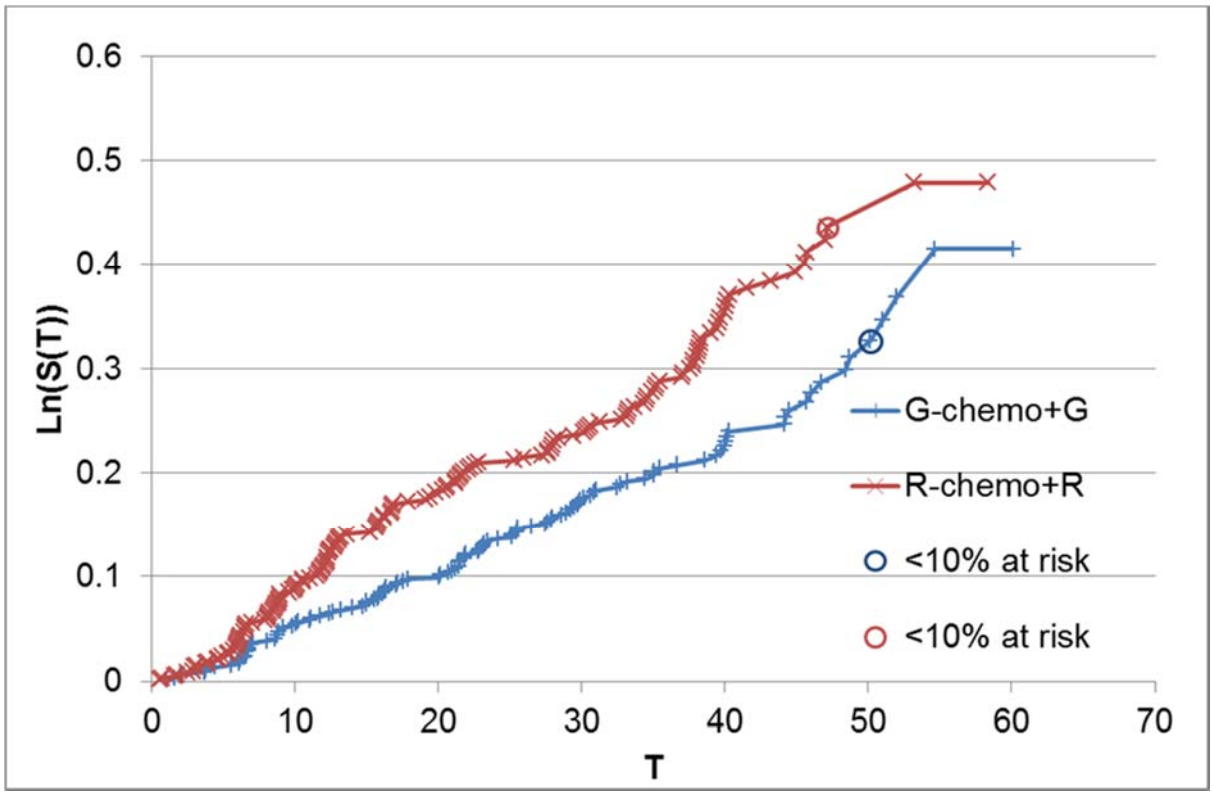
Visual inspection of the log-cumulative hazards plots for PFS in the R-chemo+R and G-chemo+G arm of GALLIUM in Figure 23 show that the curves seem to run parallel and therefore the proportional hazards assumption is valid.

Figure 23: Log-cumulative hazard plot for PFS in GALLIUM (ITT FL population)



For the extrapolation of PFS beyond the trial period parametric functions were therefore fitted simultaneously for both arms, G-Chemo+G and R-Chemo+R, with treatment as a covariate in the model. The visual inspection of the cumulative hazard plot in Figure 24 also supports a proportional hazard (constant factor between the curves).

Figure 24: Cumulative hazard plot GALLIUM PFS INV – FL ITT



Goodness of fit of the PFS parametric functions

Parametric distributions were fitted to the patient level data in both arms, with treatment as a covariate, and assessed for their goodness of fit to the data using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The results with the respective rank are shown in Table 63.

Table 63: Parametric functions, AIC and BIC goodness of fit for PFS

Distribution	AIC	Ranking	BIC	Ranking
EXPONENTIAL	1785.9	5	1796.1	3
WEIBULL	1782.2	4	1797.5	5
LLOGISTIC	1779.9	3	1795.1	2
LNORMAL	1774.5	1	1789.7	1
GAMMA	1776.4	2	1796.8	4
GOMPERTZ	1785.9	6	1801.2	6

AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion

Overall, Log-Normal, Log-Logistic or Gamma functions presented the best fit to the observed data according to AIC or BIC values. However, Exponential, Weibull or Gompertz presented still plausible fits to the observed GALLIUM data (Figure 25). Moreover, the quality and

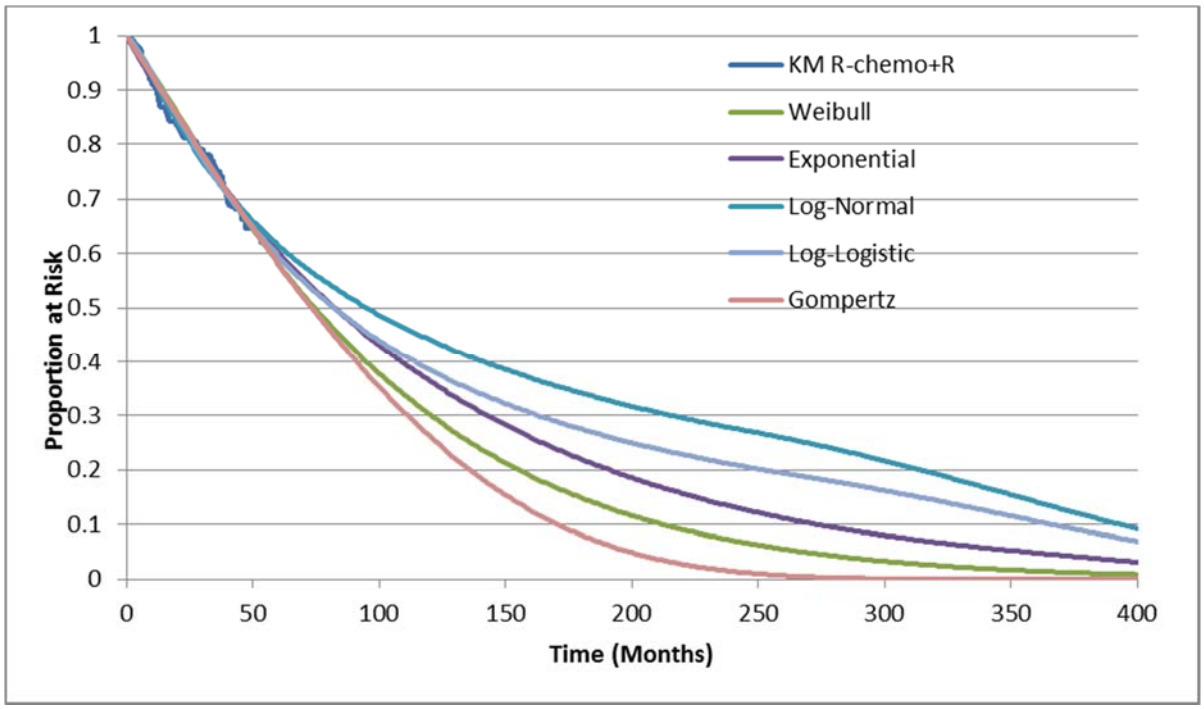
plausibility of the extrapolation beyond the observation period cannot be assessed mathematically and are therefore not reflected in AIC or BIC value. In addition to the AIC/BIC statistics, Cox-Snell residuals were used to assess the absolute fit of the models. This did not seem to favour or rule out any of the functions. In the light of the data immaturity it was therefore not feasible to rule out any of the functions based on goodness of fit to the observed data and visual inspection and external validity of the tail was therefore more important in selecting plausible functions.

Visual inspection and external validity

The overall proportion of patients remaining in PFS was restricted by mortality in PFS. This was implemented so that the risk of death or progression was always higher than the risk of death in the general UK population in the model, avoiding implausible long term PFS estimates, such as PFS curves crossing general population survival.

Figure 25 below shows the different models fitted to the R-chemo+R (i.e. the standard of care) arm in GALLIUM, all models presented plausible fits to the observed data. However, they differed in their long term predictions of PFS.

Figure 25: PFS extrapolations, R-Chemo+R arm in GALLIUM (FL ITT population)



To further select plausible forms for long-term PFS extrapolation, the predictions of the different parametric functions were compared to the observed long term behaviour in other data sets for the comparator R-chemo+R arm. These data were from the PRIMA study

(Salles et al., 2011, Seymour JF et al., 2013) and a publication from the US LymphoCare registry (Nastoupil et al., 2015).

PRIMA is the main Phase III, randomised controlled trial of MabThera maintenance in patients with high tumour burden FL responding to MabThera plus chemotherapy induction (Salles et al., 2011, Seymour JF et al., 2013). Roche had access to patient level data from the study and was able to construct an R-chemo+R arm from the data set by combining data from the induction phase with R-chemo (pre-randomisation in PRIMA) with the data from the R-maintenance arm. Details of the PRIMA study and the analysis of PFS, PPS and OS for the R-chemo+R group are described in Appendix 6. The patient characteristics of PRIMA and GALLIUM were broadly similar. However, at the time PRIMA was conducted bendamustine was not available and therefore only data for patients receiving CHOP or CVP in induction was available for comparison.

As described in Appendix 6, the follow up data was available for 8 years of an R-chemo+R cohort.

An alternative source of long term outcomes is the US LymphoCare registry. Nastoupil et al. (Nastoupil et al., 2015) report outcomes for patients enrolled in LymphoCare with stage III/IV follicular lymphoma receiving R-CHOP (n=287), R-CVP (n=187) or R with a fludarabine-based regimen (R-Flu) (n=137) as frontline therapy. Of these patients 45%, 61% and 51% received R maintenance in the follow up period for R-CHOP, R-CVP and R-Flu, respectively. The median follow up was 7.4 years. The 7 year PFS rate in R-CHOP is slightly lower than 50% (~47%) and in R-CVP is lower at 40% (Figure 2 in (Nastoupil et al., 2015)).

One of the main limitations of the LymphoCare data is that not all patients potentially eligible for maintenance may have received maintenance as the registry enrolled prior to the wider use of maintenance after first-line induction. The proportion receiving maintenance, e.g. 45% to 61% of all patients starting induction, is less than the 85% observed in PRIMA or in UK clinical practice. Similar to PRIMA, LymphoCare did not present long term follow up data on R-benda+R as bendamustine has only recently been more widely used in the first line treatment of FL. A further limitation is that LymphoCare enrolled only patients from US centres.

Long term PFS extrapolations of the different functions fitted to the GALLIUM R-chemo+R arm and the observed KM from PRIMA (R-chemo+R) and are shown in Figure 26 and PFS rates at different time points in Table 64. Within the range of observed PFS behaviour, Exponential, and Log-Logistic functions seem to predict PFS rates in the observed range.

Log-normal and Generalised Gamma would seem to predict PFS at the high end and Weibull at the lower end, respectively. Conversely, the Gompertz distribution underestimates observed PFS (also with the LymphoCare cohort in Nastoupil) and can therefore be ruled out.

Figure 26: PFS extrapolations for R-chemo+R arm

[firgure redacted]

Table 64: PFS rates at different time points for parametric functions

	PFS at 6yrs (%)	PFS at 8yrs (%)	PFS at 10yrs (%)	PFS at 15yrs (%)
Exponential	54.6	44.6	36.4	22.0
Weibull	51.3	39.6	30.2	14.9
Log-logistic	54.1	45.2	38.5	27.5
Log-normal	57.1	49.8	44.1	34.2
Generalized Gamma	56.8	49.3	43.5	33.3
Gompertz	50.8	37.4	26.2	8.1

In a UK advisory board, consultants recommended using a function representing the mid-range of plausible estimates, i.e. Exponential or Log-logistic. In the base case, an Exponential function was therefore selected. Alternative functions were investigated in insensitivity analyses. Base-case parameters for the Exponential distribution are shown below.

Table 65: PFS Base case PFS fit parameters and covariance matrix- Exponential

Fit		Covariance	
Parameter		**Intercept**	**Treatment**
Intercept	5.135	0.0083	-0.0083
Treatment (R-chemo)	-0.358	-0.0083	0.0145

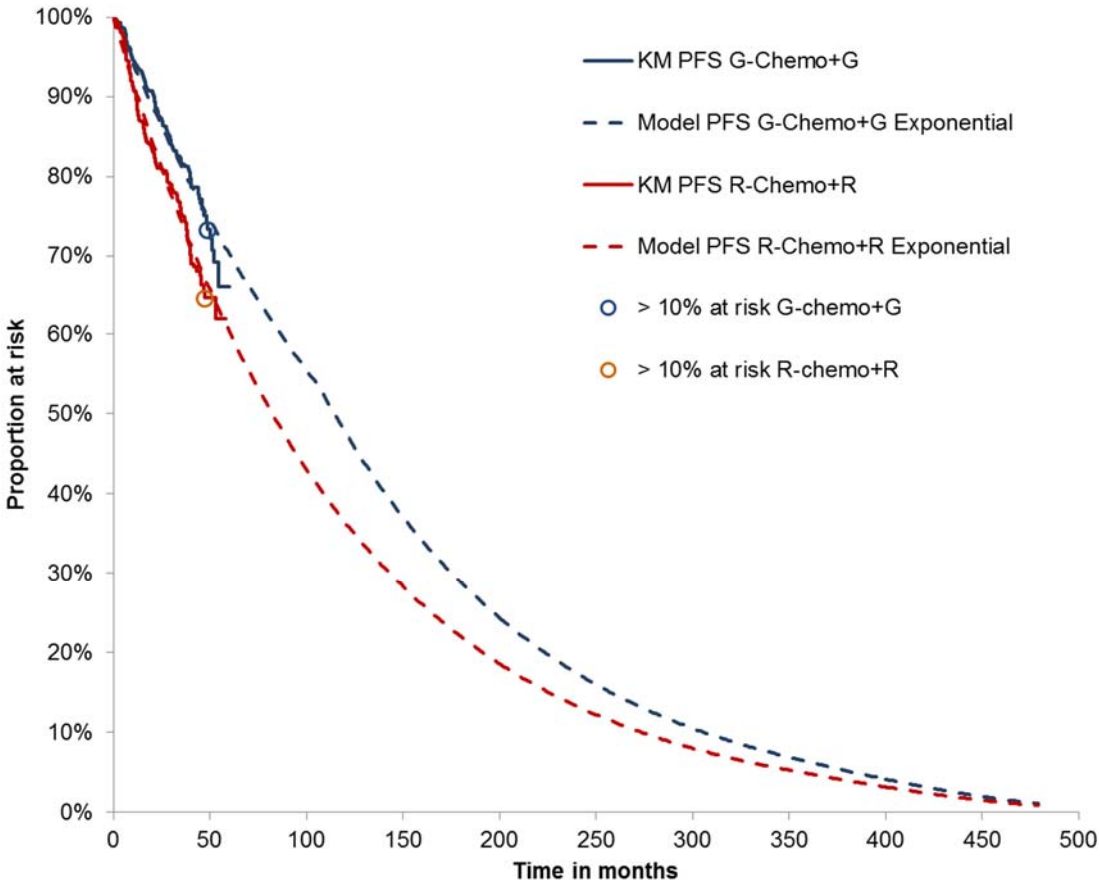
Long term PFS on G-chemo+G

To model the long term PFS on G-chemo+G a constant hazard (proportional hazard assumption) was applied. Based on the observed long term follow up in the PRIMA study there was no indication of a finite duration of treatment effect on PFS in the FL setting, i.e. the proportional hazard assumption for PFS seemed to hold for the entire observation period with longest follow up reaching of up to 9.75 years. Clinical advisors suggested that there is no evidence of a finite duration of treatment effect in treatments of FL and that it is plausible that this will be the case for G-chemo+G versus R-chemo+R.

In the model, a simple time dependent hazard was implemented to test the sensitivity of different assumptions on duration of treatment effect, i.e. to model a potential non-constant hazard in the future. Although it is not expected that the hazard function changes suddenly in reality, the constant hazard from GALLIUM was applied for a fixed period only (duration of PFS treatment effect) and a hazard of one (no treatment effect) was assumed beyond this period. In the base case, a treatment effect of 9.75 years was assumed (in line with PRIMA) and sensitivity to this parameter was tested in the sensitivity analysis. However, it is plausible that no upper limit on the duration of effect applies, e.g., Papaioannou et al. and Dewilde et al do not seem to make explicit assumptions on duration of effect for PFS. However, Papaioannou et al. did investigated scenarios where the time a patient could spend in PFS was limited in sensitivity analyses.

Figure 27 below shows the base case PFS extrapolation.

Figure 27: PFS base case extrapolation



Probability of transitioning from PFS to Death

Disease progression and mortality are competing risks for each patient in the PFS health state. In order to calculate the proportion of patients who died before progression, the model considered the UK age-specific all-cause mortality rates and the PFS death rate observed in the GALLIUM study and uses the greater value of the two rates to determine the proportion transitioning to death from PFS. In particular in the long term, mortality in PFS is expected to be driven by age related background mortality.

The probability of death in PFS was derived from the observed mortality in PFS in the GALLIUM study. Since there were few events, number of deaths and the number of patient-months at risk in PFS were pooled between the arms. The respective figures are shown below.

Table 66: Monthly death rates in PFS in GALLIUM (ITT FL)

Events	Patient months at risk	Monthly rate (95%CI)
39	39519	0.099% (0.072%- 0.135%)

Probability of death from Early and Late PD and post-progression survival

It is known that patients progressing earlier have different outcomes than those progressing later. Specifically, Casulo et al. reported from the LymphoCare study that FL patients developing an event within 24 months of diagnosis and after initial treatment with R-CHOP and R-CVP had inferior survival than those who did not have an event within 24 months (Casulo et al., 2015b). Of 588 patients treated with R-CHOP, five-year overall survival was 50% in the early-progression group compared to 90% in the reference group that did not have an event 2 years after diagnosis. Patients with early progression also had inferior post progression survival compared with those whose progression occurred after 2 years (HR 1.89; 95% CI, 1.18 to 3.03; p=0.008).

Similar findings were reported for European cohorts by Maurer et al. (Maurer et al., 2016), who showed that immunochemotherapy treated patients who relapse before 24 months had poor outcomes compared to those who did not progress within 24 months.

It should also be noted that the difference in outcomes between early and late progression could not be explained by differences in baseline FLIPI score in Casulo et al.; Maurer et al. also concluded that FLIPI was no longer prognostic in early progression and that therefore reassessment of patient status 12–24 months after diagnosis was a powerful prognostic tool in follicular lymphoma—superseding the baseline FLIPI score.

Post-progression survival data sources

To allow for differences in post progression mortality based on time to progression, the model has two progressed disease sates (PD) for early and late progression after first initial symptomatic treatment. Different post-progression survival assumptions are used in the base case for Early and Late PD, respectively. Data from the GALLIUM trial was used to inform the Early PD mortality. For Late PD, data from PRIMA was used for post-progression survival as longer term follow up was available from this study. Due to the finding by Casulo et al. and Maurer et al. PPS was not assumed to depend on baseline FLIPI. Alternative assumptions of post-progression survival for Early and Late PD were explored in sensitivity analyses.

To derive post progression mortality rates, individual patient level data on the time from progression to death was taken from the GALLIUM study. Due to the indolent nature of the disease, the data was immature and a relatively small number of events were available for analyses. The data was analysed by pooling the treatment arms and stratifying for early and late progression events. The results are shown in the figure below. As there were no PPS events in late progression only early progression PPS was used in the base case.

For the PRIMA data, the R-chemo+R cohort was analysed as described in Appendix 6.

The PPS KM curves for the early and late PD data sets from PRIMA are shown in Figure 28.

Figure 28: PPS in PRIMA for early PD (within 2 years) vs late PD (subsequent years)


[Figure redacted]

In addition, data from PRIMA was not stratified by early and late progression and a pooled rate for death in PD was derived for a scenario analyses.

Monthly mortality rates used in the base case and sensitivity analyses in the model are shown below. In the model, the greater of the UK general population and the trial cohort mortality rates in Table 67 is applied to the PD to death transition to account for the expected increase in long term mortality due to age.

Table 67: Monthly death rates in PD

	GALLIUM	PRIMA	PRIMA POOLED
Early progression (<2yrs)	1.61%	0.93%	0.77%
Late progression (>2yrs)	#N/A	0.56%	0.77%

As usual within the Markov approach, OS was an outcome of the model as the sum of time spent in PFS, early or late PD, respectively. The OS model outcomes in relation to the trial data and results from other models are discussed in Section 5.7.3.

5.3.2 Demonstrate how the transition probabilities were calculated from the clinical data. If appropriate, provide the transition matrix and describe the details of the transformation of clinical outcomes or any other relevant details here

See description in 5.3.1.

5.3.3 If there is evidence that (transition) probabilities may change over time for the treatment effect, condition or disease, confirm whether this has been included in the evaluation. If there is evidence that this is the case, but it has not been included, provide an explanation of why it has been excluded.

See Section 5.3.1. Time dependent PFS functions were implemented as described. For, post-progression mortality there was no evidence from the trial data of post progression mortality changing over time. However, the constant post progression mortality rate was compared to the age dependent general population background mortality and the greater of the values was used.

5.3.4 If clinical experts have assessed the applicability of the clinical parameters or approximated any of the clinical parameters, provide the following details:

- **the criteria for selecting the experts**
- **the number of experts approached**
- **the number of experts who participated**
- **declaration of potential conflict(s) of interest from each expert whose opinion was sought**
- **the background information provided and its consistency with all the evidence provided in the submission**
- **the method used to collect the opinions**
- **the medium used to collect opinions (for example, was information gathered by direct interview, telephone interview or self-administered questionnaire?)**

- **the questions asked**
- **whether iteration was used in the collation of opinions and if so, how it was used (for example, the Delphi technique).**

The advice on the development of this submission and economic model was sought from UK clinical experts and health economists to assess the applicability of the model inputs.

5.4 Measurement and valuation of health effects

Health-related quality-of-life data from clinical trials

5.4.1 If health-related quality-of-life (HRQL) data were collected in the clinical trials identified in section 4, comment on whether the data are consistent with the reference case. Consider the following points, but note that this list is not exhaustive:

- **method of elicitation**
- **method of valuation**
- **point when measurements were made**
- **consistency with reference case**
- **appropriateness for cost-effectiveness analysis**
- **results with confidence intervals**

Patient reported outcomes (PRO) in GALLIUM were evaluated through a validated lymphoma-specific instrument, the Functional Assessment of Cancer Therapy-Lymphoma (FACT-Lym) questionnaire and a generic, validated preference-based, health utility questionnaire, the EuroQol-5D Questionnaire (EQ-5D) as reported in section 4.7. The EQ-5D summary scores were collected at baseline, during treatment, after treatment, at the last assessment prior to progression, and at the first assessment after progression. It is important to note that the questionnaire was administered before any other study procedure was performed during the study visit. Due to the fact that PROs were only collected at the first assessment after progression, PROs were not available from GALLIUM beyond the point of progression. EQ-5D utility scores and FACT-Lym were therefore available in PFS (covering induction, maintenance and observation) and in progression at first assessment after progression was detected. The main limitation of the collected EQ-5D utility scores in GALLIUM is therefore the lack of long-term data on patients beyond progression.

The EQ-5D health index showed no statistically significant overall difference between the G-chemo+G and R-chemo+R arms over time during the treatment and follow-up periods.

To inform the health state utilities in the economic model and to compare GALLIUM data to EQ-5D values to the literature, 5,007 observations from 1,097 patients were analysed with a mixed effects model with health states in Table 68 as categorical effect, and the following baseline covariates: centralis ed age, baseline utility, ECOG, gender and FLIPI score.

Table 68: GALLIUM EQ-5D utility scores and covariance matrix

Health State	LSM estimate for the utility	Covariance					
		Induction - off tx	Induction - on tx	Maintenance & follow-up - off tx	Maintenance & follow-up - on tx	Early progression ≤ 2yrs	Late progression > 2yrs
Induction - off tx	0.772	0.0004	0.0002	0.0002	0.0002	0.0002	0.0002
Induction - on tx	0.823	0.0002	0.0006	0.0002	0.0002	0.0002	0.0002
Maintenance & follow-up - off tx	0.818	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
Maintenance & follow-up - on tx	0.831	0.0002	0.0002	0.0002	0.0006	0.0002	0.0002
Early progression ≤ 2yrs	0.776	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
Late progression > 2yrs	0.814	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002

The comparison of the EQ-5D values reported in GALLIUM with utility values in FL from the literature is discussed in section 5.4.5.

5.4.2 If applicable, describe the mapping methods used to estimate health state utility values from the quality-of-life data collected in clinical trials. Please include the following information:

- **which tool was mapped from and onto which other tool (for example, SF–36 to EQ–5D)**
- **details of the methodology used**
- **details of validation of the mapping technique**
- **if the mapping technique is published or has been used in other NICE technology appraisals for similar diseases or health conditions.**

EQ-5D values measured directly in GALLIUM were available. Mapping was therefore not required. Mapping functions to EQ-5D utility values for the lymphoma specific FACT-Lym instrument collected in addition to EQ-5D in GALLIUM are not available.

Health-related quality-of-life studies

5.4.3 Describe how systematic searches for relevant HRQL data were done. Consider published and unpublished studies, including any original research commissioned for the technology. Provide the rationale for terms used in the search strategy and any inclusion and exclusion criteria used. The search strategy used should be provided in an appendix

Search strategy development

The systematic review of utility studies was developed for utilities in previously untreated or treated follicular lymphoma. Details of the search filters are shown in Appendix 5.

Data sources

Electronic databases searched included MEDLINE, MEDLINE In-Process, EMBASE and NHS EED. Hand searches were conducted in abstracts including the following organisations:

- American Society of Hematology (ASH)
- American Society of Clinical Oncology (ASCO)
- European Society of Medical Oncology (ESMO)
- European Haematology Association (EHA)
- International Conference on Malignant Lymphoma (ICML)
- International Network of Agencies for Health Technology Assessment (INAHTA)
- National Institute for Health and Care Excellence (NICE)
- National Institute for Health Research HTA
- Canada: Canadian Agency for Drugs and Technologies in Health (CADTH)

Search implementation date and span

Initial searches were conducted on 4th April 2017 and included publications from1998 onwards, coinciding with the introduction of rituximab. Hand searches covered the years from 2012 onwards and were conducted in April 2017.

Study selection process

Screening

Studies were screened for inclusion in the full text stage by two independent reviewers using an adaptation of the PICOS framework in Table 69 with the respectiv inclusion/exclusion criteria.

Inclusion and exclusion criteria

The inclusion criteria were defined using an adaptation of the PICOS framework in Table 69 and applied at screening and the full text review stage as below.

Table 69: PICOS for utility review

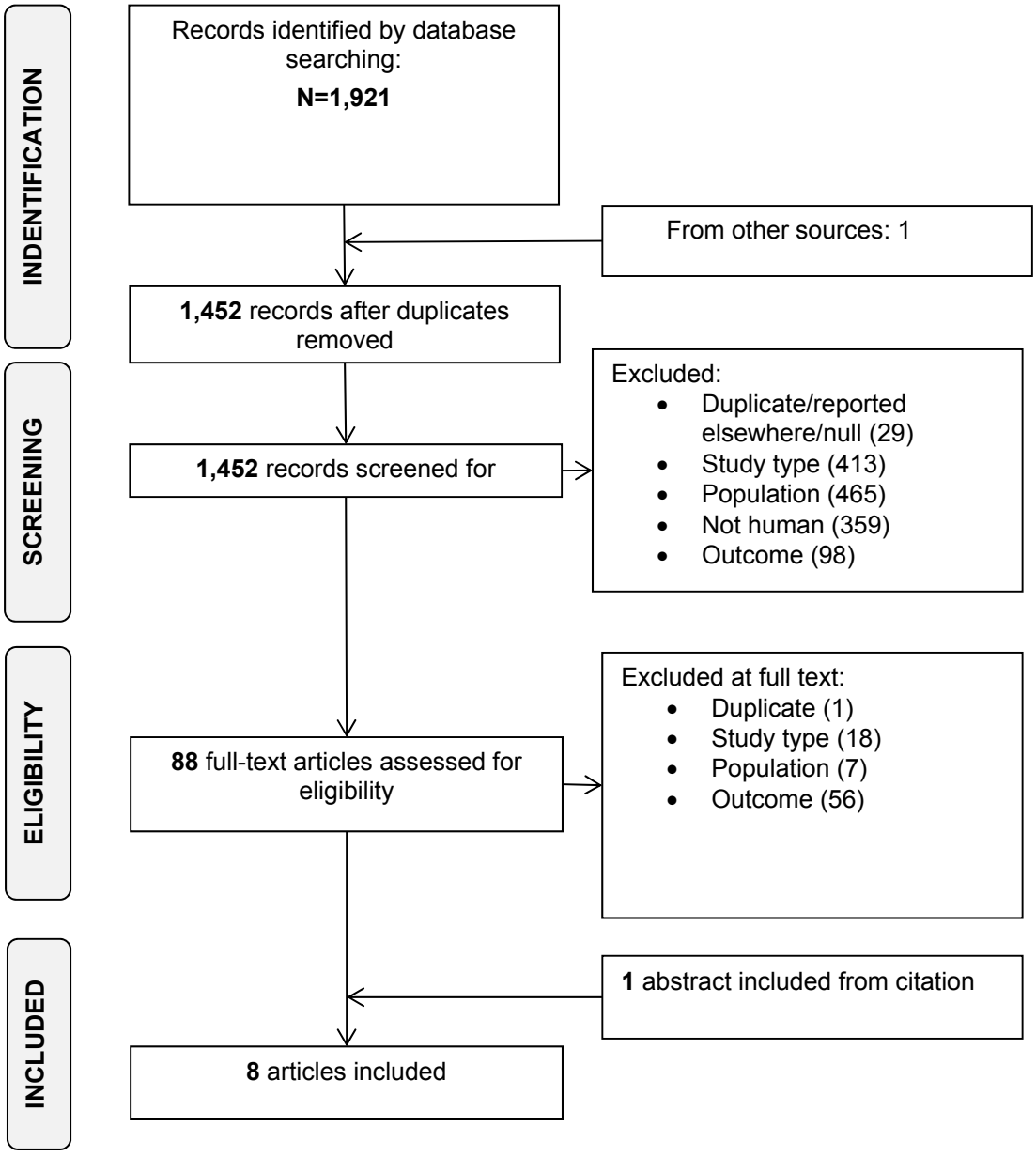
PICOS	Inclusion criteria	Exclusion criteria
Population	People with previously treated or untreated iNHL (in particular FL) All subtypes, except skin lymphomas	• Disease area not iNHL
Intervention and Comparators	• Not restricted, any intervention	
Outcomes	• Utility, preference-based HRQOL measures (e.g. EQ-5D)	• Outcomes not of interest: i.e. non generic preference-based HRQOL measures
Study types	• Studies for utility instruments • Clinical trials reporting outcome of interest	• Study design or publication format not of interest, including: o Secondary publications o Review articles, systematic literature reviews or meta-analyses o Editorials or notes or letters to the editor o Studies containing no primary data.

Results

The PRISMA flow-diagram outlining the study selection process is presented in Figure 29.

88 references were reviewed in full text. Of these, 81 were excluded due to duplication (1), study type (18), population (7) or outcome (55). In addition, one abstract was included based on citation as a primary source for utilities used in economic evaluations identified during the full text review. Eight studies were therefore included in the narrative review of utility studies in iNHL.

Figure 29: PRISMA diagram for identification of utility studies



Details of included studies

5.4.4 Tabulate the details of the studies in which HRQL was measured. Include the following, but note that this list is not exhaustive:

- **population in which health effects were measured**
- **information on recruitment (for example, participants of a clinical trial, approximations from clinical experts, utility elicitation exercises including members of the general public or patients)**
- **interventions and comparators**

- **sample size**
- **response rates**
- **description of health states**
- **adverse reactions**
- **appropriateness of health states given the condition and treatment pathway**
- **method of elicitation**
- **method of valuation**
- **mapping**
- **uncertainty around values**
- **consistency with reference case**
- **appropriateness for cost-effectiveness analysis**
- **results with confidence intervals**
- **Appropriateness of the study for cost-effectiveness analysis.**

The eight identified studies are summarised in Table 70 below. The studies were further reviewed for their appropriateness to inform health state utilities in the model and only two studies Wild D et al. (Wild D et al., 2006) and Bec M et al. (Bec M et al., 2014) were potentially suitable.

Wild D et al. collected data on 222 patients with FL in eight centres in the UK. Utilities were elicited from patients using the EQ-5D questionnaire and clinical data collected allowed allocation of patients to 5 health states:

- Active disease, newly diagnosed
- Active disease, relapsed
- Partial response to therapy
- Complete response to therapy/remission
- Disease free (no detectable disease)

Measurements were also pooled to derive utilities for pre-progression, i.e. PFS, and post-progression (PD health state reported in Table 70).

Bec M et al. report EQ-5D scores in a cross-sectional study of iNHL patients across Europe collected in an on-line questionnaire. The study included data from 18 UK patients and reported utility values for PFS and PD.

Therefore, a limited number of studies were identified that could inform utility values for the model. The most relevant literature utility values to inform the PFS and PD disease states in this model is the cross-sectional study by Wild D et al., which reported the largest sample of UK FL patients. However, the utilities from all references had some key limitations for the appropriate use within the model:

- Studies lacked a distinction between PFS utility whilst on treatment and PFS utility whilst off treatment. Although the study of 222 UK patients reported a difference HRQoL in the FACT-LYM score between people receiving chemotherapy versus those who did not (Pettengell et al., 2008), this did not translate into a significant difference in EQ-5D utility value.
- Studies reported limited data on utility depending on line or treatment, e.g. first or subsequent progression or remission and were limited by sample size.

Due to the indolent nature of FL, studies lacked longitudinal follow up to study utility with disease course and age.

Table 70: Studies included in the utility review

Title/Author	Intervention and comparators	Population and sample size	Instrument/ Method of valuation	Method of elicitation	Mapped to	HRQoL values	Original source	Appropriateness for use in model
Evaluating treatment strategies in advanced Waldenström macroglobulinemia: use of quality-adjusted survival analysis. (Levy et al., 2001)	Fludarabine vs cyclophosphamide, doxorubicin and prednisone	Waldenström macroglobulinemia	Unclear	Unclear	N/A	Unclear	Unclear	Not applicable, patients with Waldenström macroglobulinemia only.
Utility Elicitation in Patients with FL. (Wild D et al., 2006)	N/A	Patients with FL (n=222)	EQ-5D	Survey	N/A	PFS: 0.805 (SE: 0.018) PD: 0.618 (SE: 0.056)	N/A	Appropriate for model: • PFS and PD reported in a large sample of 222 UK patients with FL
Discrimination of health states in follicular lymphoma with utilities derived from the EuroQOL EQ5D instrument. (Friedlich et al., 2006)	N/A	Patients FL or other iNHL (n=84)	EQ-5D	Survey	N/A	• All: 0.84 (+/- 0.24). • Observation: (0.91 +/- 0.16) • First remission: (0.84 +/- 0.25) • Subsequent remissions: (0.81 +/- 0.20) • Active chemotherapy: (0.75 +/- 0.27). • Ongoing remission: (0.88 +/- 0.21) Not in remission (0.80 +/- 0.22)	N/A	Not appropriate for model: • Small, single centre in Canada Unclear health state definition
Determinants of the optimal first-line therapy for	R-CHOP, R-Flu and R-CVP	Advanced FL	Extrapolation from literature	Model	N/A	• Receiving RCHOP 0.70 • Receiving RFlu 0.75		Not appropriate: Unclear how

follicular lymphoma: A decision analysis. (Olin et al., 2010)			values			• Receiving RCVP 0.85 • Remission 0.99 • Remission with prolonged cytopenias 0.90		derived
Outcome and quality of life favour a conservative treatment of patients with primary gastric lymphoma. (Fischbach et al., 2011)	N/R	49 diagnosed MZL , MALT and DLBCL		Survey	N/A	N/A	N/A	Not appropriate: No FL patients
Psychosocial factors associated with impact of cancer in long-term haematological cancer survivors (Korszun et al., 2014)	N/R	718 long-term haematological cancer survivors in London	EQ-5D	Survey	N/A	Aggregate Utility not reported	N/A	Not applicable: Utilities not reported by health state and disease.
Long-term efficacy of 90Y ibritumomab tiuxetan therapy in follicular non-Hodgkin lymphoma and health-related quality of life. (Andrade-Campos et al., 2014)	90Y ibritumomab tiuxetan	Patients with FL in single Spanish centre	SF36	Survey	N/A	HRQoL z-score (SD) of SF-36 Spanish population: - Physical functioning: -0.09 - Physical role: -1.25 - Bodily pain: -0.29 - Physical health: 0.33 - Vitality: 0.39 - Social functioning: 0.03 - Emotional role: -1.8 - Mental health: -0.48	N/A	Not applicable: Utilities not mapped, single Spanish centre.
French Utility Elicitation in Previously Treated Patients	N/A	Patients with previously treated iNHL	EQ-5D	Web survey questionnaire	N/A	Reported per country for PFS and PD:	N/A	Appropriate for model, with

with Indolent Non-Hodgkin Lymphoma (iNHL). (Bec M et al., 2014)		Germany (n=5) Italy (n=18) UK (n=18) Spain (n=18) France (n=16)				• France: 0.68 (0.59-0.77) and 0.45 (0.30-0.60) • Germany: 0.84 (0.77-0.90) and 0.66 (0.55-0.78) • UK: 0.71 (0.63-0.79) and 0.51 (0.37-0.64) • Spain: 0.74 (0.65-0.83) and 0.53 (0.39-0.68) Italy: 0.82 (0.76-0.88) and 0.71 (0.62-0.81)		limitations: • PFS and PD utilities reported Small number (n=18) of UK patients

5.4.5 Highlight any key differences between the values derived from the literature search and those reported in or mapped from the clinical trials

Compared to studies identified in the literature, GALLIUM utilities were based on a very large sample of 1,097 patients (5,007 observations) with previously untreated FL. GALLIUM EQ-5D utility values in PFS appear in general higher than those reported in the cross-sectional sample in Wild D et al.2006 and significantly higher than those reported in the study by Bec M et al. (Table 71). One explanation may be that GALLIUM only captured previously untreated patients whereas the cross-sectional studies may have included pre-treated patients. EQ-5D scores collected in GALLIUM after progression seem also considerably higher than EQ-5D utility values reported for patients classified as 'progressive disease' in both the Wild D et al. and Bec M et al studies. Again, this could be due to the fact that the cross-sectional studies were not focused on previously untreated patients. In addition, a limitation of the GALLIUM data with respect to progressed disease is that it did not capture advanced stages of progression as data was only collected at first assessment after progression or at the visit that resulted in an detection of progression and not beyond.

Table 71: Utility values from GALLIUM and literature

	Mean utility value (Standard Error)		
Health State	**GALLIUM**	**Wild 2006**	**Bec UK sample**
PFS (Induction - off tx)	0.772	0.81 (0.02)	0.71 (0.04)
PFS (Induction - on tx)	0.823		
PFS (Maintenance & follow-up - off tx)	0.818		
PFS (Maintenance & follow-up - on tx)	0.831		
Early progression (≤ 2yrs)	0.776	0.62 (0.06)	0.51 (0.07)
Late progression > 2yrs	0.814		

Adverse reactions

5.4.6 Describe how adverse reactions affect HRQL. The effect of adverse reactions on HRQL should be explored regardless of whether they are included in a cost-effectiveness analysis in the base-case analysis. Any exclusion of the effect of adverse reactions on HRQL in the cost-effectiveness analysis should be fully justified.

Disutilities for adverse events (AEs) were not included in the base-case. AEs were similar between the two treatment arms and including disutilities for AEs in the model in a sensitivity analysis did not result in a significant effect on the overall QALYs in each arm and the incremental difference between arms. This is also supported by the fact that EQ-5D utility

scores values were similar for patients on- and off-treatment in GALLIUM and the effects of AEs while on treatment may have been captured in the collected utility values (Table 68).

In a sensitivity analysis, disutilities for AEs from literature sources were applied for AEs of Grade three and above that occurred in more than 2% of patients, according to the values in Table 72.

Table 72: AR disutilities for sensitivity analysis

Grade 3/4 adverse event	Disutility	SE	Source	Duration of adverse event (days)	Source
Neutropenia	-0.09	0.02	(Nafees et al., 2008)	15.10	NICE TA 306 2013
Thrombocytopenia	-0.11	0.02*	(Tolley et al., 2013)	23.20	NICE TA 306 2013
Anaemia	-0.12	0.02	(Swinburn et al., 2010)	16.07	NICE TA 306 2013
Leukopenia	-0.12	0.02	Assumed to be same as Anaemia	16.07	-
Pneumonia	-0.20	0.02	(Beusterien et al., 2010)	14.00	NICE TA 306 2013

*SE (standard error) is assumed to be the average of all other AE disutility standard errors

Health-related quality-of-life data used in cost-effectiveness analysis

5.4.7 Define what a patient experiences in the health states in terms of HRQL in the cost-effectiveness analysis. Explain how this relates to the aspects of the disease or condition that most affect patients' quality of life.

In PFS, on-treatment induction, patients are expected to typically respond to treatment, with associated improvements in HRQoL. On the other hand AEs associated with chemotherapy may reduce HRQoL.

In PFS, induction off-treatment, patients are expected to have a reduced HRQoL compared to patients on treatment because they are expected to be off-treatment due AEs or non-response to treatment.

In PFS, on maintenance, patients are expected to have responded to induction treatment, with associated improvements in HRQoL. Conversely, AEs associated with ongoing maintenance may reduce HRQoL.

PFS (maintenance & follow up) off treatment are expected to have responded to induction treatment and completed maintenance resulting in a good HRQoL.

In early PD patients are expected to progress quickly and require further treatment for symptomatic disease, with a reduced HRQoL compared to the PFS state.

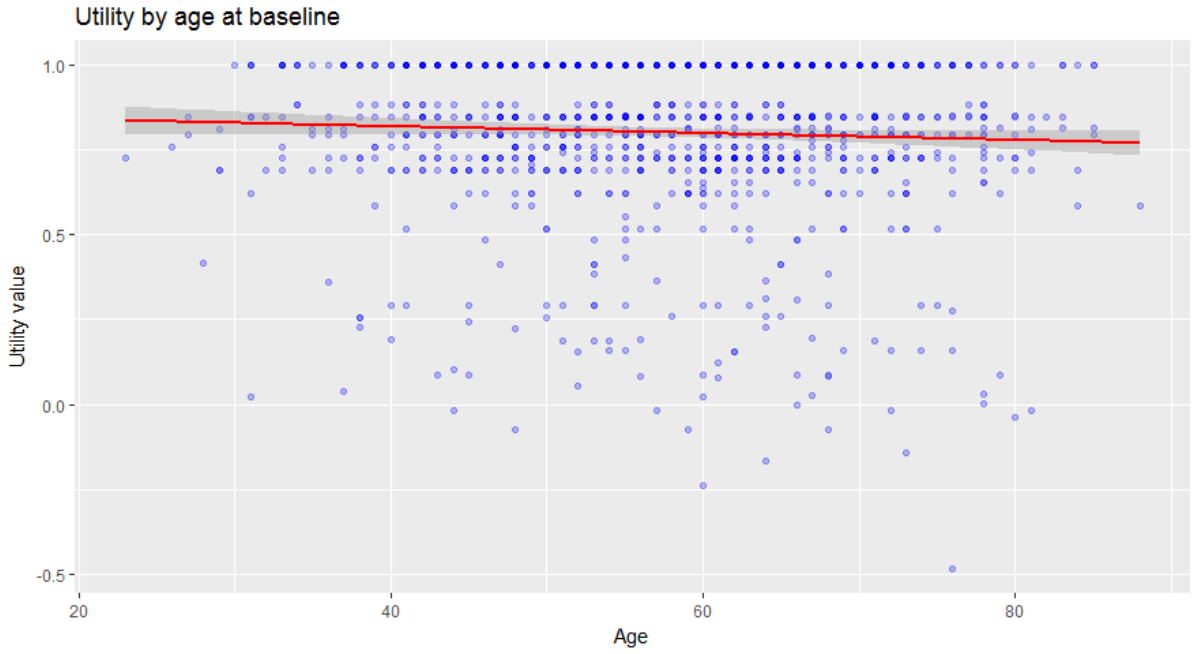
In late PD patients are expected to progress and require further treatment for symptomatic disease, with a reduced HRQoL compared to the PFS state. However, disease progression is expected to be slower than for early PD.

Patient experience is described in section 3.2.

5.4.8 Clarify whether HRQL is assumed to be constant over time in the cost-effectiveness analysis. If not, provide details of how HRQL changes over the course of the disease or condition.

EQ-5D baseline values collected in GALLIUM at baseline (Figure 30) did not show an age dependent decline. Therefore, utility values in the model were not adjusted for age in the base case.

Figure 30: GALLIUM baseline EQ-5D scores by age



In a sensitivity analysis, health state utilities in PFS and PD in the model utilities were adjusted for age effects seen in the general UK population (Ara and Brazier, 2011) by applying an age-related multiplier to reduce otherwise constant utilities with age in each health state.

5.4.9 If appropriate, describe whether the baseline HRQL assumed in the cost-effectiveness analysis is different from the utility values used for each of the health states. State whether quality-of-life events were taken from this baseline.

No quality-of-life events were taken from baseline utility values in PFS or PD.

5.4.10 If the health state utility values used in the cost-effectiveness analysis have been adjusted, describe how and why they have been adjusted, including the methodologies used.

No adjustments were made.

5.4.11 Identify any health effects found in the literature or clinical trials that were excluded from the cost effectiveness analysis and explain their exclusion.

No health effects associated with FL were excluded.

5.4.12 In a table, summarise the utility values chosen for the cost-effectiveness analysis, referencing values obtained in sections 5.4.1–5.4.6. Justify the choice of utility values, giving consideration to the reference case. For continuous variables, mean values should be presented and used in the analyses. For all variables, measures of precision should be detailed. See below for a suggested table format.

Base-case utility values for PFS were from the GALLIUM study as EQ-%D HRQoL was directy reported from patients with a large representative sample in PFS. For the PD states in the model, values from Wild D et al. were used as summarised in Table 73. The main reason for this was that utility values for PD were deemed more representative for the advanced stages of the disease not captured at the point of first progression in the GALLIUM study. In sensitivity analyses, the use of EQ-5D utility scores at progression from GALLIUM (Table 68) were explored.

Table 73: Summary of utility values for cost-effectiveness analysis

Health state	Utility value: mean	Standard Error*	Reference in submission (section)	Justification
PFS (Induction - off tx)	0.772	0.027	Section 5.4.1	GALLIUM trial based estimates
PFS (Induction - on tx)	0.823	0.007		
PFS (Maintenance & follow-up - off tx)	0.818	0.005		
PFS (Maintenance & follow-up - on tx)	0.831	0.006		
Early PD (including subsequent treatments)	0.62	0.06	Section 5.4.3	Value from Wild D et al. representative of later disease stage captured in the model progressed disease states.
Late PD (including subsequent treatments)	0.62	0.06	Section 5.4.3	
AEs	See Table 72		Section 5.4.6	

*See covariance matrix Table 68

5.5 Cost and healthcare resource use identification, measurement and valuation

Resource identification, measurement and valuation studies

5.5.1 All parameters used to estimate cost effectiveness should be presented clearly in a table with details of data sources. For continuous variables, mean values should be presented and used in the analyses. For all variables, measures of precision should be detailed.

5.5.2 Describe how relevant cost and healthcare resource use data for England were identified. Include the search strategy and inclusion criteria, and consider published and unpublished studies to demonstrate how relevant cost and healthcare resource

use data for England were identified. The search strategy used should be provided in an appendix. If the systematic search yields limited

- **data for England, the search strategy may be extended to capture data from**
- **other countries. Please give the following details of included studies:**
- **country of study**
- **date of study**
- **applicability to clinical practice in England**
- **cost valuations used in the study**
- **costs for use in the economic analysis**
- **technology costs.**

Search strategy development

The search strategy was developed using a combination of free text, MEDLINE MeSH and EMBASE EMTREE terms, as appropriate for the databases included. Briefly, the search terms in the strategy included:

- Disease state terms for previously untreated iNHL
- Cost and resource use.

Details are shown in Appendix 5.

Data sources

The electronic databases EMBASE, MEDLINE, MEDLINE In-Process and other non-indexed citations (NHS EED and EconLit) were searched.

To ensure the most recent published data are included, the following congresses were also searched:

- International Society for Pharmacoeconomics and Outcomes Research (ISPOR)
    - European and North American congresses
- European Society for Medical Oncology (ESMO)
- American Society of Clinical Oncology (ASCO)
- American Society of Hematology (ASH)
- European Haematology Association (EHA)

In addition, relevant economic evaluations were reviewed to determine the source of UK resource data being used.

Search implementation date and span

Electronic databases were searched from 1st January 1998 to 13th March 2017. Congresses were hand searched from 2014 to March 2017.

Study selection process

After duplicates were removed, titles and abstracts were reviewed by two independent reviewers. A third reviewer arbitrated any differences against the eligibility criteria. Full text articles were then reviewed against the inclusion and exclusion criteria.

Inclusion criteria

The inclusion criteria are outlined using the PICOS framework in Table 74.

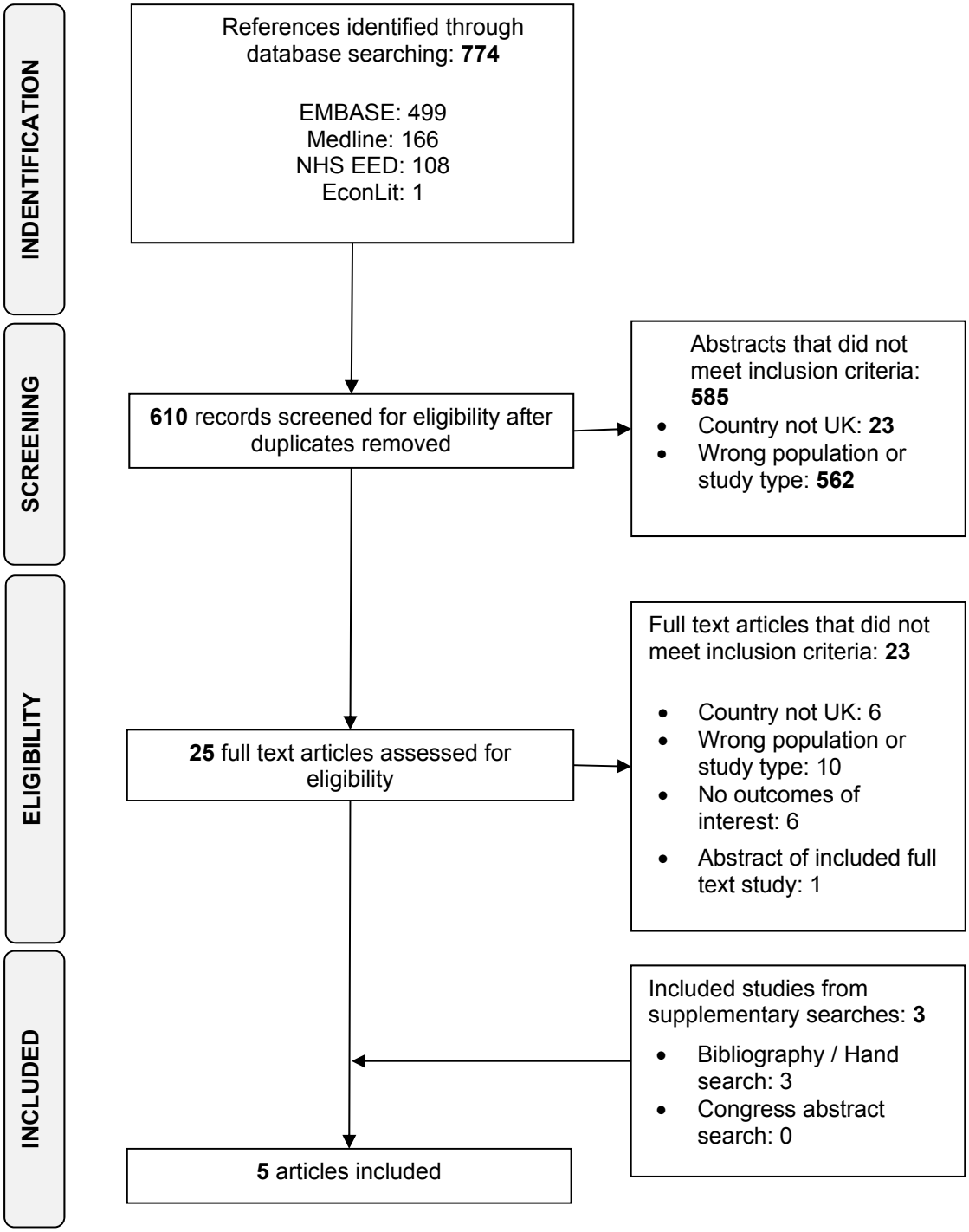
Table 74: PICOS of the resource use SLR

	Inclusion	Exclusion
Population	• Adults with treatment-naïve / previously untreated iNHL; in particular FL • Preference for studies with a UK perspective	• Animal/in vitro studies • Children (≤18 years old) • Non-UK studies • Patients with lymphoma cell types other than iNHL (e.g. aggressive NHL such as DLBCL), skin lymphoma iNHL subtype, or mantle cell lymphoma
Intervention & Comparator	• All approved or investigational pharmacotherapies • Palliative/Supportive care • No treatment being investigated	• Alternative medicine (such as homeopathy, naturopathy, and Reiki)
Outcomes (to be included if at least one is reported)	• All cost-related outcomes • All medical resource use outcomes	
Study design	• Including, but not limited to studies of: • Resource use • Cost analyses • Economic burden	
Limits	• English language only	
Timespan	1998 – present (start date chosen as this was the launch of rituximab and reflects current standard of care)	

Results of the resource use SLR

Electronic searches were performed on 13th March 2017. After removing duplicates, 610 were screened at title and abstract level excluding 585 references. Following full text review of the remaining 25 references, two studies were included and three further studies were included based on hand searches. The PRISMA flow diagram is shown below.

Figure 31: PRISMA diagram for the resource use SLR



Summary of identified studies

All included studies were conducted from the perspective of the UK NHS and PSS in patients with indolent NHL or FL receiving first-line treatment. Four were comparative economic evaluations reporting both costs and clinical outcomes and one was a model developed to predict lifetime costs. Publication years ranged from 2006 to 2016. Details of each included study are reported in Table 75.

Lewis et al., 2006 (Lewis G et al., 2006) reported a Markov model developed to evaluate the cost-effectiveness of rituximab combined with cyclophosphamide, vincristine, and prednisone (CVP) compared with CVP alone. This study was published as an abstract only and reported lifetime healthcare costs per patient but did not provide details of resource use or cost inputs and sources.

Ray et al. 2010 conducted a cost-effectiveness analysis of rituximab compared with commonly used chemotherapy regimens for patients with advanced FL. The model was published as a full paper and drug and health state resource use, costs and sources were described in detail.

As part of the re-review of TA243 the External Review Group (ERG) undertook a de novo cost-effectiveness analysis of rituximab compared with commonly used chemotherapy regimens for patients with advanced FL reported in Papaioannou et al., 2012. Later, Dewilde et al., 2014 developed a model based on the methodology presented by Papaioannou et al., 2012 to compare R-benda(+R) to R-CHOP(+R) in the first line setting.

Wang et al, 2016 developed a discrete event simulation model to estimate cost of current care. The model was published as a poster at ISPOR 2016. Costs and life expectancy are presented and described briefly.

Table 75: Studies reporting UK health care resource utilisation in FL

Author and date	Lewis *et al.*, 2006	Ray *et al.*, 2010	Papaioannou *et al.*, 2012	Dewilde *et al.*, 2014	Wang *et al.*, 2016
Country and perspective	UK NHS & PSS	UK NHS & PSS	UK NHS & PSS	UK NHS & PSS	UK NHS & PSS
Population	Follicular NHL	Follicular NHL	Follicular NHL	Indolent NHL	Follicular NHL
Treatment status	First-line	First-line	First-line	First-line	First, second and third-line
Comparators	CVP R-CVP	CVP R-CVP CHOP R-CHOP MCP R-MCP CHVP + IFNa (1st 6 months) R-CHVP + IFNa (1st 6 months) CHVP + IFNa (2nd 6 months) R-CHVP + IFNa (2nd 6 months)	CVP R-CVP CHOP R-CHOP MCP R-MCP	R-B R-CVP CHOP R-CHOP FC R-FC HDT R-HDT GCSF Stem cell ASCT	Chemo R-Chemo Radiotherapy ASCT
Reported resource use					
Drug acquisition	X (limited data)	X	X	X	X (limited data)
Drug administration		X	X	X	
Side effects			X	X	
PFS health state		X	X	X	
PD health state		X			
Palliative care			X	X	
Death			X	X	
Total lifetime costs	X				X
Cost per QALY	X				
Applicability	Directly	Directly	Directly	Directly	Partially

5.5.3 When describing how relevant unit costs were identified, comment on whether NHS reference costs or payment-by-results (PbR) tariffs are appropriate for costing the intervention being appraised. Describe how the clinical management of the condition is currently costed in the NHS in terms of reference costs and the PbR tariff. Provide the relevant Healthcare Resource Groups and PbR codes and justify their selection with reference to section 2.

In line with recent health economic evaluations for the first line treatment of FL identified in the literature (Dewilde et al., 2014, Papaioannou et al., 2012), resource for administration and supportive care were based on NHS reference costs and PSSRU unit costs as described in the sections below.

5.5.4 If clinical experts assessed the applicability of the cost and healthcare resource use values available, or approximated any of the values used in the cost-effectiveness analysis, provide the details (see section 5.3.4).

See section 5.3.4.

Intervention and comparators' costs and resource use

5.5.5 In a table, summarise the cost and associated healthcare resource use of each treatment. A suggested format for a table is provided below. Cross refer to other sections of the submission; for example, drugs costs should be cross-referenced to section 2.3.1. Provide a rationale for the choice of values used in the cost-effectiveness model discussed in section 5.2.2.

A summary of the unit costs associated with acquisition and administration of the intervention and comparator medicines is presented in Table 76 with further details described below.

Table 76: Unit costs associated with the technology

Items	Intervention (G-chemo+G)	Comparator (R-chemo+R)	Reference in submission
Gazyvaro 1st cycle induction	Cycle 1: £9,936.00 (PAS [redacted])	-	Table 77
Gazyvaro subsequent induction cycles and maintenance	£3,312.00 (PAS [redacted])	-	
MabThera IV per cycle induction or maintenance	-	£1218.04 ([redacted])	
MabThera SC (maintenance)	-	£1344.65 ([redacted])	Table 77
Bendamustine per cycle induction	£91.73		

Cyclophosphamide per cycle induction	£12.37		
Doxorubicin per cycle induction	£7.51		
Vincristine per cycle induction	£6.28		
Prednisolone per cycle induction	£3.10		
Administration, Pharmacy & patient transport 1st cycle bendamustine induction	£1583.32	£814.66	Table 79
Administration, Pharmacy & patient transport 1st cycle CVP or CHOP induction	£1198.99	£430.33	
Administration, Pharmacy & patient transport - Subsequent bendamustine induction	£814.66		
Administration, Pharmacy & patient transport - Subsequent CVP or CHOP induction	£430.33		
Administration, Pharmacy & patient transport - maintenance IV	£354.55		
Administration, Pharmacy & patient transport - maintenance SC	-	£264.83	

Drug acquisition costs

Drug costs were from the British National Formulary (British National Formulary, 2017) or eMIT (Department of Health, 2016), where available. For all drugs, actual average doses used in the GALLIUM study were used in the model base case.

Details of the drug cost calculations are described below and are summarised in Table 77.

Table 77: Drug acquisition costs

Drug	Cost per vial	Cost per cycle*	Dosing
Gazyvaro	£3,312 per 1,000 mg	Cycle 1: £9,936.00 Subsequent cycles: £3,312.00 Maintenance: £3,312.00	Cycle 1: 3 fixed dose of 1,000 mg (days 1, 8 and 15) Subsequent cycles: one fixed dose of 1,000 mg Maintenance (for responders): one fixed dose of 1000 mg every 2 months for up to two years or until progression.
Gazyvaro (induction) with PAS	[redacted] with PAS per 1,000 mg	Cycle 1: [redacted] Subsequent cycles: [redacted] Maintenance: [redacted]	
MabThera IV	100 mg vial at £174.63 and 500 mg vial at £873.15	£1218.04 (based on 1.86 BSA)	Per cycle: 375 mg/m^2 Maintenance (for responders): 375 mg/m^2 every 2 months for up to two years or until progression.
MabThera IV [redacted]	100 mg vial at [redacted] and	[redacted] (based on 1.86 BSA)	

	500 mg vial at [redacted]		
MabThera SC List	£1344.65 per 1,400 mg	£1344.65	Maintenance (for responders): 1,400 mg every 2 months for up to two years or until progression.
MabThera SC ([redacted])	[redacted] per 1,400 mg	[redacted]	
Bendamustine	£27.77per 100 mg vial, £6.85 per 25 mg vial	£91.73 (based on 1.86m^2)	Cycles 1–6: 90 mg/m^2/day on days 1 and 2.
Cyclophosphamide	500 mg vial at £7.84, 1000 mg vial at 8.87	£12.37 (based on 750 mg/m^2 and 1.86m^2)	Local protocols
Doxorubicin	50 mg at £4.04	£7.51 (based on 50 mg/m^2 and 1.86m^2)	Local protocols
Vincristine	1 mg at £3.14	£6.28 (based on 2 mg)	Local protocols
Prednisolone	30 5 mg tablets at £0.93	£3.10 (based on 500 mg)	Local protocols

*Model uses actual doses per cycle from GALLIUM

Gazyvaro induction with benda, CHOP or CVP

In the G-chemo arm, eight to ten doses of Gazyvaro at 1000 mg were administered by IV infusion with the accompanying chemotherapy regimen during induction.

- G-CHOP: Gazyvaro was administered on Days 1, 8, and 15 of Cycle 1 and on Day 1 of Cycles 2–8 (21-day cycles). CHOP was administered on Day 1, with prednisone/prednisolone/methylprednisolone also administered on Days 2–5 of Cycles 1–6.
- G-CVP: Gazyvaro was administered on Days 1, 8, and 15 of Cycle 1 and on Day 1 of Cycles 2-8 (21-day cycles). CVP was administered on Day 1, with prednisone/prednisolone/methylprednisolone also administered on Days 2–5 of Cycles 1–8.
- G-bendamustine: Gazyvaro was administered on Days 1, 8, and 15 of Cycle 1 and on Day 1 of Cycles 2-6 (28-day cycles). Bendamustine was administered on Days 1 and 2 of Cycles 1-6, with prednisone/prednisolone/methylprednisolone administered on Day 1 of Cycle 1.

The list price for a 1,000 mg vial of Gazyvaro is £3,312.00 [redacted] with PAS). This equates to a cost of £9,936.00 [redacted] with PAS) in cycle 1 and £3,312 [redacted] with PAS) in subsequent cycles.

Gazyvaro maintenance

For patients responding to induction therapy Gazyvaro is administered by intravenous infusion at a fixed 1,000 mg dose once every 2 months for up to 2 years or until disease progression at a cost of £3,312 [redacted] with PAS).

MabThera induction

In the R-chemo arm, six to eight doses of MabThera at 375 mg/m^2 were administered by IV infusion with the accompanying chemotherapy regimen during induction.

- R-CHOP: MabThera was administered on Day 1 of Cycles 1-8 (21-day cycles). CHOP was administered on Day 1, with prednisone/prednisolone/methylprednisolone also administered on Days 2–5, of Cycles 1-6.
- R-CVP: MabThera was administered on Day 1 of Cycles 1–8 (21-day cycles). CVP was administered on Day 1, with prednisone/prednisolone/methylprednisolone also administered on Days 2–5, of Cycles 1–8.
- R-bendamustine: MabThera was administered on Day 1 of Cycles 1-6 (28-day cycles). Bendamustine was administered on Days 1 and 2 of Cycles 1-6, with prednisone/prednisolone/methylprednisolone also administered on Day 1 of Cycle 1.

BNF list price for MabThera is 100 mg vial at £174.63 and 500 mg vial at £873.15.

[redacted]

MabThera maintenance

For patients responding to induction therapy MabThera is administered by intravenous infusion MabThera at 375 mg/m^2 with 100 mg vial at £174.63 and 500 mg vial at £873.15. However, MabThera is also used as subcutaneous formulation in clinical practice in England, with about [redacted]% of eligible patients receiving MabThera SC in maintenance (Roche Products Ltd.). The model uses the MabThera SC costs for one 1,400 mg vial at the nationally available net price of [redacted] (List price £1344.65).

Chemotherapy regimens

Benda is now available as generic formulations in vials of 25 mg and 100 mg at a cost of £6.85 and £27.77 per vial, respectively. The model uses the actual dose used which was consistent with a planned dose of 90 mg/m^2 on Days 1 and 2 of each cycle.

CHOP and CVP regimens were administered with G or R according to the standard preparation and infusion procedures of each investigational site. The model uses the average actual dose in the GALLIUM study. Unit costs were based on eMIT data (version May 2016) for Cyclophosphamide (500 mg vial at £7.84, 1000 mg vial at 8.87), Doxorubicin (50 mg at £4.04), Vincristine (1 mg at £3.14), Prednisolone (30 5 mg tablets at £0.93). Actual average doses reported in GALLIUM were consistent with a typical dosing (see e.g. http://www.londoncanceralliance.nhs.uk/information-for-healthcare-professionals/forms-and-guidelines/south-east-london-cancer-network/haematology/non-hodgkins-lymphoma/) of cyclophosphamide at 750 mg/m^2, doxorubicin at 50 mg/m^2 and vincristine at 1.4 mg/m^2 (up to 2.0 mg), and prednisolone at five times 100 mg per cycle.

Drug administration costs

Drug administration costs in the model are based on NHS references costs tariffs (NHS Schedule of Reference Costs 2016 (Department of Health, 2017). Additional pharmacy costs for the preparation of the infusion and patient transport costs were included.

The cost of £407 (SB14Z - Deliver Complex Chemotherapy, Prolonged infusion, at First Attendance – Day case) is applied for each first attendance of each cycle and the cost of £361 (SB15Z - Deliver subsequent elements of a chemotherapy cycle) for each subsequent attendance. The cost of £337 (SB 13Z, Deliver more Complex Parenteral Chemotherapy at First Attendance) was applied for administration of MabThera IV or Gazyvaro as maintenance and £253 (SB12Z, Deliver Simple Parenteral Chemotherapy at First Attendance, day-case) was applied for administration of MabThera SC as maintenance.

Pharmacy costs were based on 15 minute preparation time (Papaioannou et al., 2012) and £46 per hour hospital pharmacist (band 6) unit costs of £11.50 (Curtis, 2016). SC administration was assumed to require no pharmacy costs. With Papaioannou et al. 30% of patients were assumed to require NHS transportation at a cost of £39.24, i.e. an average transportation cost of £11.77 per administration was assumed.

The applicable administration costs in the R-chemo+R and G-chemo+G schedule are summarised below.

Table 78: Administration schedule and applicable costs for R-chemo+R and G-chemo+G

Applicable cost	Induction						Maintenance
	Cycle 1				Subsequent cycles		
	Day 1	Day 2	Day 8	Day 15	Day 1	Day 2	Day 1
SB12Z (£253)							R SC only
SB13Z (£337)							IV only
SB14Z (£407)	X				X		
SB15Z (£361)		Benda only	G only	G only		Benda only	
Pharmacy preparation costs (£11.50)	X	Benda only	G only	G only	X	Benda only	IV only
NHS transportation (£11.77)	X	Benda only	G only	G only	X	Benda only	X

Administration costs per cycle for all combinations of anti-CD20, chemotherapy, and induction or maintenance, are summarised in Table 79.

Table 79: Administration costs per cycle

Scenario	Tariff	Pharmacy	Transport	Total
1st Cycle G-benda+G	£1490.00	£46.00	£47.08	£1583.09
1st Cycle G-CHOP+G, 1st Cycle G-CVP+G	£1129.00	£34.50	£35.31	£1198.81
1st and subsequent cycles R-benda+R	£768.00	£23.00	£23.54	£814.54
1st and subsequent cycles R-CHOP+R, and R-CVP+R	£407	£11.50	£11.77	£430.27
G or R IV maintenance cycle	£337	£11.50	£11.77	£360.27
R SC maintenance cycle	£253	-	£11.77	£264.77

Administration costs per maintenance cycle based on the values above, was therefore 25% lower for MabThera SC compared to MabThera IV. In the model, this reduction is applied in proportion to the patients receiving MabThera SC in maintenance.

Health-state unit costs and resource use

5.5.6 Summarise and tabulate the costs included in each health state. A suggested format for a table is provided below. Cross refer to other sections of the submission

for the resource costs. Provide a rationale for the choice of values used in the cost-effectiveness model. The health states should refer to the states in section 5.2.2.

Drug acquisition and administration costs are described in section 5.5.5. These are converted into monthly costs, depending on the cycle length, and applied to the PFS (on-treatment) health state. Supportive care costs in PFS (on and off treatment) and PD and costs for subsequent treatments in PD after progression are described in detail below. Table 80 below contains a summary of all health state costs in the model.

Table 80: Summary of health state costs

Health states	Items	Average cost per patient		Reference in submission
		Intervention (G-chemo+G)	Comparator (R-chemo+R)	
PFS (on treatment)	Acquisition costs (list) Cycle 1 antiCD20 (G or R)	£9,936.00	£1218.04 (IV)	Table 77
	Acquisition costs (list) Subsequent Cycles antiCD20 (G or R)	£3,312.00	£1218.04 (IV)	
	Acquisition costs (list) maintenance Cycles antiCD20 (G or R)	£3,312.00	£1218.04 (IV) £1344.65 (SC)	
	Acquisition costs chemo per cycle	£91.73 (benda) £21.75 (CVP) £29.26 (CHOP)		
	Administration costs (cycle 1)	£1490.00 (w. benda), £1129.00 (w. CHOP or CVP)	£1490.00 (w. benda), £1129.00 (w. CHOP or CVP)	Table 79
	Pharmacy costs (cycle 1)	£46.00 (benda), £34.50 (CHOP, CVP)	£46.00 (benda), £34.50 (CHOP, CVP)	
	Transport cost (cycle 1)	£47.08 (benda), £35.31 (CHOP, CVP)	£47.08 (benda), £35.31 (CHOP, CVP)	
	Administration costs subsequent cycles	£768.00 (benda), £407 (CHOP, CVP)		
	Pharmacy costs subsequent cycles	£23.00 (benda), £11.50 (CHOP, CVP)		
	Transport, subsequent cycles	£23.66 (benda), £11.77 (CVP, CHOP)		
	Administration costs maintenance	£337	£337	
	Pharmacy costs maintenance	£11.50 (IV only)		
	Transport, maintenance	£11.77		

PFS (on and off treatment)	Supportive care induction (initial 6 months or progression)	£253.27	Table 81
	Follow up supportive care 6 to 30 months or progression	£82.59	
	Follow up supportive care beyond 30 months or progression	£57.82	
Early PD	Supportive care	£231.27	
	Subsequent treatment and supportive care	£13,427	Section 'Subsequent treatment costs in PD'
Late PD	Supportive care	£57.82	Table 81
	Subsequent treatment and supportive care	£13,427	Section 'Subsequent treatment costs in PD'

PFS and PD health state supportive care cost

In the absence of UK data or guidelines, it was assumed that patients have initially monthly haematologist visits during induction therapy (accounted for in the first six months in the model). The frequency of follow up visits was then assumed to be initially three months as stated in ESMO guidelines (Dreyling et al., 2016). The frequency of visits for patients remaining progression-free is then assumed to decrease to four visits per month as suggested in the ESMO guidelines and Papaioannou D et al. Costs for the visit were based on NHS reference cost and laboratory costs reported in Papaioannou D et al. It was also assumed that patients receive one CT scan during induction (0–6 months) and then in the follow up period (6–30 months).

In progressed disease, patients may receive subsequent lines of treatment and more intensive follow up. It was assumed that patients progressing early would have monthly haematologist visits and diagnostic test and examinations, whereas patients that progress late would on average require less intensive follow up. More or less intensive follow up in either late or early PD states was investigated in a sensitivity analysis.

Based on the assumed frequency, monthly supportive care costs were applied in the model. Resource costs assumed and monthly costs are summarised below.

Table 81: Supportive care costs in PFS and PD

Resource	Unit costs	Source	Frequency	Average monthly cost
PFS: induction (0–6) months				
Haematologist	£166 (£111-£209)	NHS reference costs 2015/16 Code: 303	Monthly	£166.00
Diagnostic tests/examinations*	£65.27	Sum of test costs in Papaioannou D 2012* inflated to 2015/16 prices	Monthly	£65.27
CT scan	£132 (£89-£162)	NHS reference cost 2015/16 RD27Z	Once in 6 months	£22.00
			Total	**£253.27**
PFS: follow-up 6–30 months				
Haematologist	£166.00	NHS reference costs 2015/16 Code: 303	Every three months	£55.33
Diagnostic tests/examinations*	£65.27	Sum of test costs in Papaioannou D 2012* inflated to 2015/16 prices	Every three months	£21.76
CT scan	£132.00	NHS reference cost 2015/16 RD27Z	Once in 24 months	£5.50
			Total	**£82.59**
PFS: follow-up 30 months until progression				
Haematologist	£166.00	NHS reference costs 2015/156Code: 303	Every four months	£41.50
Diagnostic tests/examinations*	£65.27	Sum of test costs in Papaioannou D 2012* inflated to 2016/16 prices	Every four months	£16.32
			Total	**£57.82**
Early PD				
Haematologist	£166.00	NHS reference costs 2015/16 Code: 303	Monthly	£166.00
Diagnostic tests/examinations	£65.27	Sum of test costs in Papaioannou D 2012* inflated to 2015/16 prices.	Monthly	£65.27
			Total	**£231.27**
Late PD				
Haematologist	£166.00	NHS reference costs 2015/156Code: 303	Every four months	£41.50
Diagnostic tests/examinations	£65.27	Sum of test costs in Papaioannou D 2012* inflated to 2016/16 prices	Every four months	£16.32
			Total	**£57.82**

*Includes: full blood count, patient history/physical examination, full profile (U&E, LFT, calcium), Serum IgG, IgA, IgM and electrophoresis & lactate dehydrogenase test.

Subsequent treatment costs in PD

Subsequent treatments were included in the model as an average cost for subsequent treatments in the early and late PD states. Clinical advisors suggested that next line treatment choices post progression would be the same between both arms and that costs and outcomes would therefore be similar. Clinical advice also suggested that treatment for early and late progressors would not differ significantly, with potentially more early progressors being considered for transplant. Although time to next anti-lymphoma treatment (NALT) was recorded in GALLIUM, this data was immature and heavily censored: patients that had progressed were censored before they received their next treatment and the follow up period was not long enough to capture higher lines of treatment. Therefore literature values were used for the subsequent treatment costs in the model. Papaioannou et al. report for the R-CHOP+R arm (Appendix 16 in Papaioannou et al.) average discounted costs of £11,795 for second-line and £1,632 for higher lines of treatments, respectively. The total of these costs of £13,427 were used in both arms of the model for early and late progression. As the model applied one costs at progression, discounted values were appropriate to account for the fact that treatments, in particular for higher lines, would occur significantly later than first progression. However, these costs are probably conservative as they were based on the average costs for all patients and not only those progressing.

In sensitivity analysis costs based on time to next anti-lymphoma treatment data from GALLIUM was used and these were calculated in Appendix 7 and resulted in a significantly lower average costs of £5,437.61 due to censoring.

Adverse event unit costs and resource use

5.5.7 Summarise and tabulate the costs for each adverse reaction listed in section 4.12 and included in the de novo cost-effectiveness analysis. These should include the costs of therapies identified in section 2.3. A suggested format for a table is provided below. Cross refer to other sections of the submission for the resource costs.

All AEs of Grades 3, 4 or 5 occurring in more than 2% of patients in either arm of the GALLIUM trial were incorporated into the model. AEs were assumed to occur at a constant rate while on treatment. The event rate for each AE was calculated as the number of observed events divided by the total patient months of exposure. The monthly rates were converted to probabilities and multiplied by the event unit costs (Table 82) to calculate average monthly AE costs for each arm and by chemotherapy strata. The average costs per month per arm were weighted by the number of patients (safety evaluable) in each

chemotherapy strata. This resulted in an average monthly AE cost while on treatment of £53.62 in the G-chemo+G arm and £45.85 R-chemo+R arm, respectively. These were applied when in PFS on treatment.

Table 82: Adverse event costs included in the model

Event (Grade)	Unit Cost	Reference
Anemia (3)	£2,117	SA03G (NL)
Febrile Neutropenia (3)	£6226.29	NICE CG NHL, 2016
Dyspnea (3)	£0.00	Not costed
Infusion related reaction (3)	£600.65	SA31E (NS)
Infusion related reaction (4)	£600.65	SA31E (NS)
Neutropenia (3)	£867.00	LRiG estimate rev. TA162, TA175
Neutropenia (4)	£867.00	LRiG estimate rev. TA162, TA175
Pneumonia (3)	£4154.97	DZ11P (NL)
Leukopenia (3)	£3236.25	SA31E (NL)
Leukopenia (4)	£3236.25	SA31E (NL)
Thrombocytopenia (3)	£3236.25	SA31E (NL)
Thrombocytopenia (3)	£3236.25	SA31E (NL)

*NHS reference costs 2015-16; NL, non-elective long stay; NS, non-elective short stay

Adverse event costs associated with subsequent treatment lines after progression were not included in the progressive health (PD) state and were assumed to be included in the costs of subsequent treatments.

Miscellaneous unit costs and resource use

5.5.8 Describe and tabulate any additional costs and healthcare resource use that have not been covered elsewhere (for example, costs relating to subsequent lines of therapy received after disease progression, personal and social services costs). If none, please state.

5.6 Summary of base-case de novo analysis inputs and assumptions

Summary of base-case de novo analysis inputs

5.6.1 Tabulate all variables included in the cost-effectiveness analysis, detailing the values used, range (for example, confidence interval, standard error or distribution) and source. Cross refer to other parts of the submission. Complete the table below that summarises the variables applied in the economic model.

See Table 83.

5.6.2 For the base-case de novo analysis the company should ensure that the cost-effectiveness analysis reflects the NICE reference case as closely as possible. Describe the rationale if an input chosen in the base-case de novo analysis:

- **deviates from the NICE reference case or**
- **is taken from other sources (such as the published literature) rather than data from clinical trials of the technology (when available).**

A summary table of the of base-case de novo analysis inputs is shown below.

Table 83: Summary of variables applied in the economic model

Variable	Value	Measurement of uncertainty and distribution	Source	Reference to section in submission
Demographics				
Age	57.9	Not applied	GALLIUM trial	5.2
Weight	75.7		GALLIUM trial	5.2
Height	168.3		GALLIUM trial	5.2
Model structure				
Time horizon	40 years	Not applied	NICE reference case	5.2
Discount rate for costs and outcomes	3.50%		NICE reference case	5.2
Transition probabilities				
Monthly probability of death from PFS	0.099%	Log-Normal	GALLIUM trial	5.3.1
Probability to remain in PFS				
PFS Exponential distribution (G-chemo+G arm, R-chemo+R) see Table 65		Covariance matrix	GALLIUM trial	5.3.1
PFS duration of treatment effect	9 years	Not applied	Assumption	5.3.1
Early PPS Lambda (λ)	1.61%	Covariance matrix	GALLIUM trial	5.3.1
Late PPS Lambda (λ)	0.56%		PRIMA trial	5.3.1
Utilities				
PFS off txt – induction	0.772	Covariance matrix	GALLIUM trial EQ5-D	5.4
PFS off txt – maint. & follow up	0.818		GALLIUM trial EQ5-D	
PFS on txt – induction	0.823		GALLIUM trial EQ5-D	
PFS on txt – maint.	0.831		GALLIUM trial EQ5-D	
Early PD	0.62	Standard Error, Beta	Wild et al. 2006	
Late PD	0.62	Standard Error, Beta	Wild et al. 2006	
Cost and resource use				
Administration costs				
1st administration in cycle	£407	25% of mean, Log-normal	SB14Z (NHS reference costs 2015-16)	5.5.5

Subsequent administrations in cycle	£361	25% of mean, Log-normal	SB15Z (NHS reference costs 2015-16)	
Maintenance administration (IV)	£337	25% of mean, Log-normal	SB13Z (NHS reference costs 2015-16)	
Maintenance administration (SC)	£253	25% of mean, Log-normal	SB12Z (NHS reference costs 2015-16)	
Pharmacy cost	£11.50	25% of mean, Log-normal	15 min PSSRU 2016	
Patient transport costs	£11.77	25% of mean, Log-normal	Papaioannou et al. 2012	
Proportion receiving SC as maintenance	[redacted]%	Not applied	Roche data on file	
Drug acquisition costs				
Gazyvaro 1,000 mg*	£3,312.00	Not applied	BNF 2017	5.5.5
MabThera IV 100 mg*	£174.63		BNF 2017	5.5.5
MabThera IV 500 mg*	£873.15		BNF 2017	5.5.5
MabThera SC 1400 mg*	£1344.65		MIMS 2017	
Bendamustine 25 mg	£6.85		BNF 2017	5.5.5
Bendamustine 100 mg	£27.77		BNF 2017	5.5.5
Cyclophosphamide 500 mg	£7.84		EMIT 2016	5.5.5
Cyclophosphamide 1000 mg	£8.87		EMIT 2016	5.5.5
Doxorubicin 50 mg	£4.04		EMIT 2016	5.5.5
Vincristine 1 mg	£3.14		EMIT 2016	5.5.5
Prednisolone, 30 5 mg tablets	£0.93	Not applied	EMIT 2016	5.5.5
Supportive care costs PFS/PD				
Haematologist visit	£166	Log-normal	NHS reference costs 2015-16 Code: 303	5.5.6
Diagnostic tests/examinations	£65.27		Papaioannou D 2012	5.5.6
CT scan	£132		NHS reference cost 2015-16 (RD27Z)	5.5.6
Subsequent treatments	£13,427	20% of mean, Log-normal	Papaioannou D 2012	5.5.6
Adverse Events				
Anaemia (3)	£2,117	Log-normal	NHS Reference Costs SA03G (NL)	5.5.7

Febrile Neutropenia (3)	£6226.29		NICE CG NHL, 2016	5.5.7
Dyspnoea (3)	£0.00		Not costed	5.5.7
Infusion related reaction (3)	£600.65		NHS Reference Costs SA31E (NS)	5.5.7
Infusion related reaction (4)	£600.65		NHS Reference Costs SA31E (NS)	5.5.7
Neutropenia (3)	£867.00		LRiG estimate rev. TA162, TA175	5.5.7
Neutropenia (4)	£867.00		LRiG estimate rev. TA162, TA175	5.5.7
Pneumonia (3)	£4154.97		NHS Reference Costs DZ11P (NL)	5.5.7
Leukopenia (3)	£3236.25		NHS Reference Costs SA31E (NL)	5.5.7
Leukopenia (4)	£3236.25		NHS Reference Costs SA31E (NL)	5.5.7

*List prices, for confidential net prices see section 5.5.5

Assumptions

5.6.3 Provide a list of all assumptions used in the de novo economic model and justify each assumption

A number of assumptions are required to make modelling this disease area feasible and to model beyond the existing data. A list of the key assumptions made when constructing this model can be found below in Table 84.

Table 84: Key assumptions in the model

Variable	Assumption	Justification/notes
Utilities	Age adjusted constant health state utilities in PFS (on/off treatment), Early and Late PD.	Studies (where available) did not show significant trends over time within health states. Constant utilities were therefore assumed that were age adjusted.
Costs and resource use	Supportive care costs in PFS or PD were assumed based on frequency of visits reported in an economic analysis in FL (Papaioannou D 2012; review of TA110) or European guidelines.	There were no studies identified that reported directly measured supportive care resources applicable to the model.
	Subsequent treatment costs in post progression were assumed to be independent of treatment arm and not significantly different in Early and Late PD.	To be consistent with assumption of treatment and outcomes in post progression being independent of treatment arm (see below).
Transitions	Progression free survival is extrapolated with an Exponential function.	Based on plausible long-term behaviour on R-chemo+R.
	A proportional hazard between intervention and comparator arm for is applied for 9 years.	Proportional hazard was observed in GALLIUM study and no finite duration of PFS treatment effect seems to have been observed in previous first line studies in FL, sucah as PRIMA.
	Post progression survival depends on time to first progression resulting in different post progression mortality in Early and Late PD.	Based on evidence in the literature and analysis of GALLIUM and PRIMA post-progression survival data.
	Post progression survival independent of treatment arm in Early and Late PD.	Patients are expected to follow a similar pathway post progression irrespective of antiCD20 treatment (R or G). There was no statistically significant difference in the observed mortality in PFS or PD between the arms in GALLIUM as the event rate was low and pooling the arms allowed a more robust estimate of the mortality rate.

5.7 Base-case results

5.7.1 Provide the results of the analysis. In particular, results should include, but are not limited to, the following:

- **the link between clinical- and cost-effectiveness results**
- **costs, quality-adjusted life years (QALYs) and incremental cost per QALY**
- **disaggregated results such as life years gained, costs associated with treatment,**
- **costs associated with adverse reactions, and costs associated with follow-up or subsequent treatment.**

Base-case incremental cost effectiveness analysis results

5.7.2 When presenting the results of the base case incremental cost effectiveness analysis in the table below, list the interventions and comparator(s) from least to most expensive. Present incremental cost-effectiveness ratios (ICERs) compared with baseline (usually standard care) and then incremental analysis ranking technologies in terms of dominance and extended dominance. If the company has formally agreed a patient access scheme with the Department of Health, present the results of the base-case incremental cost-effectiveness analysis with the patient access scheme.

The cost-effectiveness results are presented Table 85 below [redacted]

Table 85: Deterministic base case results

Technologies	Total Costs (£)	Total LYG	Total QALYs	Inc Costs (£)	Inc LYG	Inc QALYs	ICER (£/QALY)
G-chemo+G	[redacted]	13.25	9.96				
R-chemo+R	[redacted]	12.42	9.19	[redacted]	0.83	0.77	[redacted]

Values in the table are discounted and half cycle corrected

Clinical outcomes from the model

5.7.3 For the outcomes highlighted in the decision problem (see section 3), provide the corresponding outcomes from the model and compare them with clinically important outcomes such as those reported in clinical trials, as suggested in the table below. Discuss reasons for any differences between the modelled results in the cost-effectiveness analysis and the observed results in the clinical trials (for example, adjustment for crossover).

Model based results for PFS and OS are shown in Figure 32 and Figure 33, respectively. As discussed in section 5.3.1, PFS was modelled based on the actual patient level trial data in GALLIUM and therefore fitted the observed clinical trial data very well. The predicted PFS was also in the range observed in long term follow up data for the R-chemo+R arm from other data sources, for example from PRIMA, as discussed in section 5.3.1. In addition, median PFS in the standard of care arm was in line with expectations of clinical experts consulted. However, due to the indolent nature of FL, GALLIUM data did not reach median PFS, and neither did the PRIMA R-chemo+R cohort at 9 year follow up (see Appendix 6).

Overall, the predicted OS behaviour seemed plausible and in agreement with observation in GALLIUM. The model seemed to reproduce the observed OS curve in the G-chemo+G arm of GALLIUM but appeared to overestimate (until about 40 months) OS in the R-chemo+R comparator arm. However, due to the indolent nature of the disease, OS data in GALLIUM was very immature, as is generally the case in the first-line FL setting. As such, data to validate the long-term OS predictions of the model was not available. Clinical experts consulted stated that median predicted OS of around 16 years for the SOC arm exceeded their expectations of about 14 years somewhat. However, experts acknowledged that their current experience with long term survivors was based on a cohort that started treatment more than 10 years ago and that the current standard of care may potentially result in a higher OS.

Figure 32: Model base case PFS and OS (FL ITT population)

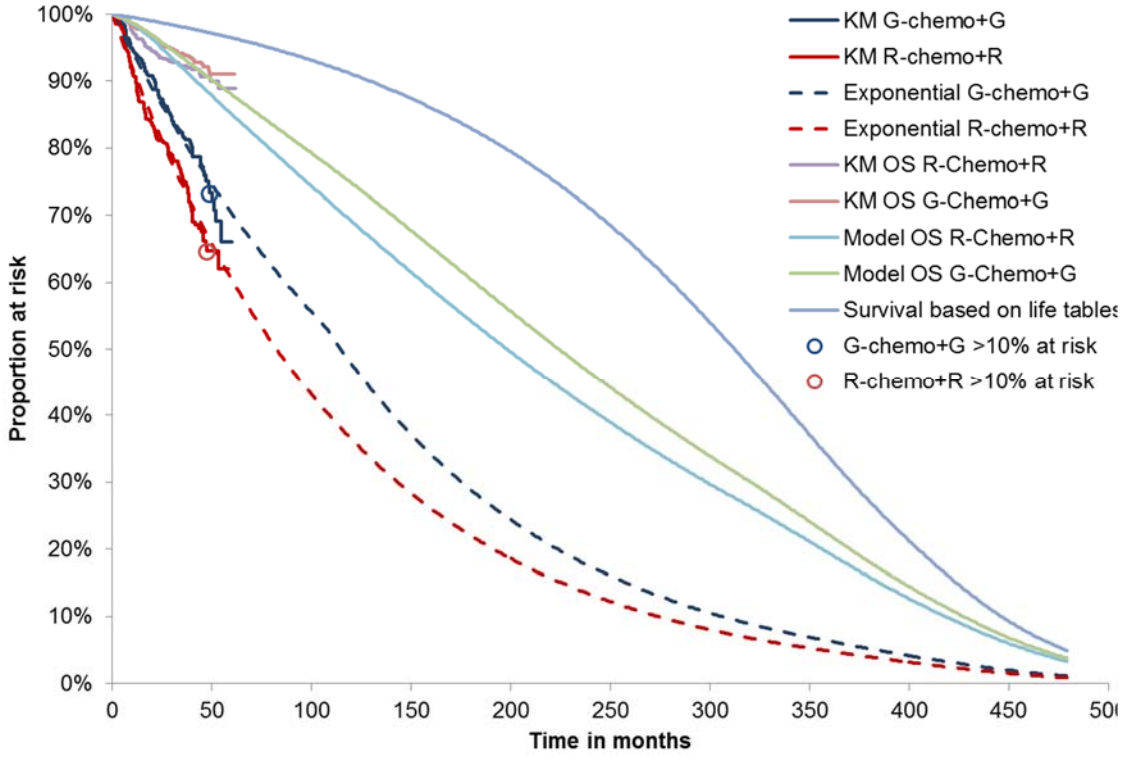
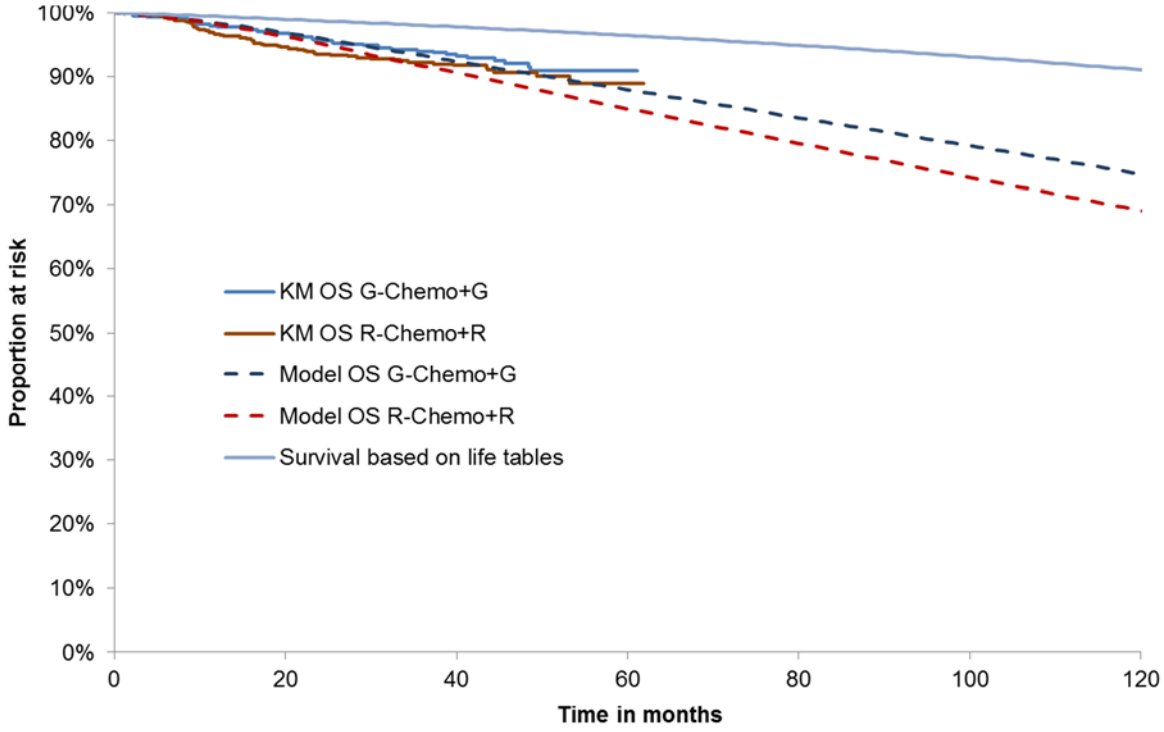


Figure 33: Model OS (FL ITT population)



The key model predictions for the base case are summarised in Table 86.

Table 86: Base case model PFS and OS outcomes (FL ITT, undiscounted)

	G-chemo+G	R-chemo+R	Difference
Mean LY in PFS	11.60	9.68	1.92
Median PFS	9.58	6.83	2.75
Total Mean LY (OS)	19.42	17.97	1.45
Median OS	18.67	16.50	2.17

Model predictions for the R-chemo+R standard of care can also be compared to predictions from models developed for rituximab in this setting. For example, Papaioannou et al. or Dewilde et al. predict mean times in PFS for R-chemo+R from 5.2 years (R-CVP, Dewilde) to 8.5 years (R-benda+R, Dewilde) and mean OS from 11.7 years (R-CVP+R; Dewilde) to 13.1 years (R-benda+R, Dewilde).

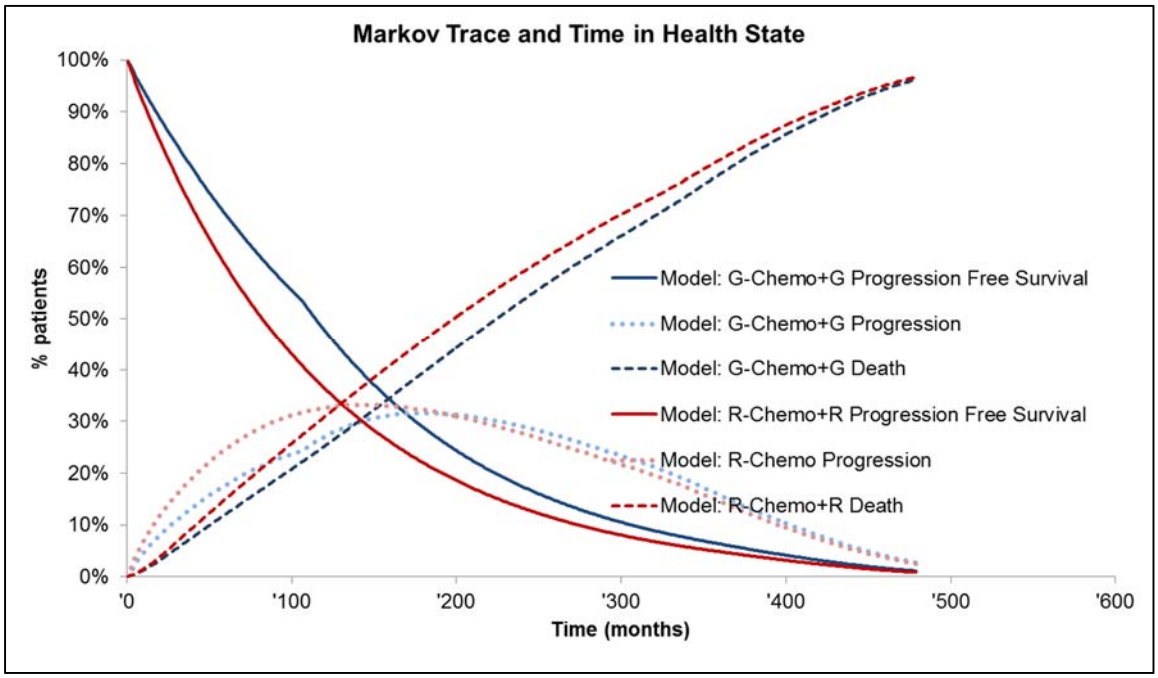
Whereas the PFS predictions are slightly lower for R-chemo+R in these models, the current analysis predicts significantly higher mean OS values in the R-chemo+R arm compared to previous approaches, and, therefore a longer time in post-progression and later lines of treatment. To obtain mean OS values in the region of 13–14 years, higher post progression mortality would have to be assumed in the model. However, this would contradict the observed mortality difference between early and late progression. A more likely explanation

is that previous models underestimated the time post first progression to death, e.g. the time in PFS after second-line treatment and in progressed disease, due to the data used to model these outcomes. For example, Papaioannou et al. used data from van Oers et al. (van Oers et al., 2010) in relapsed/refractory FL to model outcomes of second-line treatment. However, in the cohort of van Oers et al., about 50% of patients enrolled had less than two years from initial diagnosis, i.e. had therefore progressed early. In addition, approximately 20% had two prior treatments. It is therefore likely that post progression survival were underestimated in previous analyses.

5.7.4 Provide (if appropriate) the proportion of the cohort in the health state over time (Markov trace) for each state, supplying 1 for each comparator.

The Markov trace by health states (PFS, PD and Death) is shown in Figure 34 below.

Figure 34: Markov trace for base case



5.7.5 Provide details of how the model assumes QALYs accrued over time. For example, Markov traces can be used to demonstrate QALYs accrued in each health state over time.

Utilities are accrued by weighting the time in each health state as described above by the respective utilities as described in section 5.4.

Disaggregated results of the base case incremental cost effectiveness analysis

5.7.6 Provide details of the disaggregated QALYs and costs by health state, and of resource use predicted by the model in the base case incremental cost effectiveness

analysis by category of cost. The tables that should be completed summarising the disaggregated results (for example, QALY gain by health state, costs by health state, predicted resource use by category of cost) are presented below.

Disaggregated QALYs per health state are summarised in Table 87. Patients spend significantly longer average time in PFS, accounting for 78% of total absolute QALYs gained and less time in early PD (10% QALY of absolute QALY gain) and late PD (12% QALY of absolute QALY gain), respectively.

Table 87: Summary of QALY gain by health state

	G-chemo+G	R-chemo+R	Difference	Absolute	% of absolute
Health state					
Progression free survival	7.19	6.12	1.07	1.07	78%
Progression < 2 yrs	0.28	0.42	-0.13	0.13	10%
Progression > 2 yrs	2.49	2.65	-0.16	0.16	12%
Total	9.96	9.19	0.77	1.36	100%

Values in the table are discounted and half cycle corrected

Disaggregated costs per health state and cost items are summarised in Table 88 below.

Table 88: Summary of predicted resource use by category of cost

State	Cost (G-chemo)	Cost (R-chemo)	Cost difference	Absolute difference	% of absolute
PFS					
Gazyvaro	[redacted]	0	[redacted]	[redacted]	[redacted]
MabThera	0	[redacted]	[redacted]	[redacted]	[redacted]
Chemotherapy	371	365	5	5	[redacted]
Drug Administration	7,751	6,589	1,162	1,162	[redacted]
Adverse Events	1,205	986	219	219	[redacted]
Supportive Care	7,755	6,817	937	937	[redacted]
PFS Total	[redacted]	[redacted]	[redacted]		
Progressive disease					
Supportive care and subsequent treatment costs	10,201	11,873	-1,672	1,672	[redacted]
Subsequent treatment costs					
Total PD & PFS	[redacted]	[redacted]	[redacted]	[redacted]	100%

Values in the table are discounted and half cycle corrected

5.8 Sensitivity analyses

Probabilistic sensitivity analysis

5.8.1 All inputs used in the analysis will be estimated with a degree of imprecision. As specified in the NICE guide to the methods of technology appraisal, probabilistic sensitivity analysis is preferred for translating the imprecision in all input variables into a measure of decision uncertainty in the cost effectiveness of the options being compared. In non-linear decision models, probabilistic methods provide the best estimates of mean costs and outcomes. The mean value, distribution around the mean, and the source and rationale for the supporting evidence should be clearly described for each parameter included in the model. The distributions for probabilistic sensitivity analysis should not be arbitrarily chosen, but should represent the available evidence on the parameter of interest, and their use should be justified.

Provide the information specified in sections 5.8.2–5.8.4.

5.8.2 The distributions and their sources for each parameter should be clearly stated if different from those presented in section 5.5, including the derivation and value of 'priors'. If any parameters or variables were omitted from the probabilistic sensitivity analysis, please provide the rationale for the omission(s).

All model variables which had a distribution assigned are presented in Table 89. Uncertainty was characterised by standard error (if available), covariance matrix or by assuming an error of 20% from the mean if statistical uncertainty was not available. Drug acquisition costs were kept fixed.

Table 89: Parameters included in the probabilistic sensitivity analysis

Parameter	Uncertainty	Distribution
Parameters for PFS G-chemo+G/R-chemo+R arms	Covariance matrix Table 65	Multivariate normal
Probability of death in PFS	Standard Error Table 65	Log-normal
Probability of death in Early PD	Covariance matrix	Multivariate normal
Probability of death in Late PD	Covariance matrix	Multivariate normal
Utilities in PFS and PD states	Standard Error	Beta
Time on treatment	KM Greenwood CI	Log-normal
Admin costs	Standard Error	Log-normal
Pharmacy costs	20% of mean	Log-normal
Adverse event cost	20% of mean	Log-normal
Number of adverse events	Standard Error	Log-normal
Supportive care costs PFS & PD and subsequent treatments	20% of mean or Standard Error	Log-normal

5.8.3 Present the incremental cost effectiveness results of a probabilistic sensitivity analysis (including 95% confidence intervals). Include scatter plots and cost-effectiveness acceptability curves showing the probability that the treatment is cost effective if the incremental cost-effectiveness ratio ICER is £20,000 to £30,000 per QALY gained. Describe how the probabilistic ICER(s) were calculated and provide the rationale.

A 1,000 iteration probabilistic sensitivity analysis was conducted in order to determine the uncertainty surrounding the base-case ICERs. The scatter plot and the corresponding cost-effectiveness acceptability curve are shown in Figure 35 and Figure 36 respectively.

Figure 35: Incremental cost and QALY PSA base case results

[Figure redacted]

Figure 36: Cost-effectiveness acceptability curve

[Figure Redacted]

This analysis indicated that G-chemo+G was more cost-effective than R-chemo+R in [redacted]% of simulations at a threshold of £30,000/QALY gained. The probabilistic base-case ICER was £[redacted]QALY, comparable to the deterministic base-case.

5.8.4 Describe and explain, if any, the variation between the incremental cost effectiveness analysis results estimated from the base-case analysis (section 5.6) and the probabilistic sensitivity analysis.

Deterministic analysis and probabilistic sensitivity analyses results are approximately comparable.

5.8.5 Identify which variables were subject to deterministic sensitivity analysis, how they were varied, and the rationale behind this. If any parameters or variables listed in section 5.6.1 were omitted from sensitivity analysis, please provide the rationale.

A deterministic sensitivity analysis was carried out on the parameters listed in Table 90 in 5.8.6 below. Continuous parameters were varied using the 10 and 90%-percentile values obtained from the probabilistic simulation as lower and upper limits, respectively. In addition to varying continuous variables categorical variables were changed: such as parametric functions for PFS; PPS source; different settings for the time-on treatment; vial sharing and administration costs. The discount rates for costs and outcomes were varied according to standard methods and the time horizon altered.

5.8.6 Results of deterministic sensitivity analysis.

Results of the deterministic sensitivity analysis are shown in Table 90 and the tornado diagram in Figure 37.

Table 90: Deterministic sensitivity analysis for base case

Parameter modified	Base value	High Value*	Low Value*	ICER High	ICER low
Utilities					
Utility in PFS - Induction - On tx	0.823	0.834	0.812	[redacted]	[redacted]
Utility in PFS - Induction - off tx	0.772	0.783	0.761	[redacted]	[redacted]
Utility in PFS - Maintenance - off tx	0.831	0.843	0.820	[redacted]	[redacted]
Utility in PFS - Maintenance - off tx	0.818	0.830	0.806	[redacted]	[redacted]
Utility in PD - Early progression ≤ 2yrs	0.618	0.693	0.547	[redacted]	[redacted]
Utility in PD - Late progression > 2yrs	0.618	0.693	0.547	[redacted]	[redacted]
Utility source PFS	GALLIUM		Wild		[redacted]
Utility source PD	Wild		GALLIUM		[redacted]
Utility age adjusted	No		Yes		[redacted]
AR Utility included	No		Yes		[redacted]
Costs					
1st administration G-chemo	430	535	347	[redacted]	[redacted]
1st administration R-chemo	430	532	356	[redacted]	[redacted]
Administration G-chemo (subsequent)	384	423	348	[redacted]	[redacted]
Administration R-chemo (subsequent)	384	421	350	[redacted]	[redacted]
Administration maintenece G	360	454	287	[redacted]	[redacted]
Administration maintenece R	303	394	238	[redacted]	[redacted]
Supportive care PFS induction	253	292	223	[redacted]	[redacted]
Supportive care PFS maintenance	83	95	72	[redacted]	[redacted]
Supportive care PFS follow up	58	67	50	[redacted]	[redacted]
AEs - G-chemo+G	54	58	51	[redacted]	[redacted]
AEs - R-chemo+R	46	50	43	[redacted]	[redacted]
Supportive care early PD	231	272	200	[redacted]	[redacted]
Supportive care late PD	58	67	50	[redacted]	[redacted]
Subsequent treatment early PD	13,427	17,038	10,406	[redacted]	[redacted]
Subsequent treatment late PD	13,427	17,065	10,445	[redacted]	[redacted]
Subsequent treatment early/late PD	13,427		5,437.61		[redacted]
Vial sharing	Yes		No		[redacted]
Time on treatment	Actual treatment duration		According to label		[redacted]
MabThera SC use	[redacted]%	80%	40%	[redacted]	[redacted]
Outcomes					
PFS Parametric distribution function	Exponential		Weibull		[redacted]
PFS Parametric distribution function	Exponential		Log-normal		[redacted]
PFS Parametric distribution function	Exponential		Generalised		[redacted]

			Gamma		
PFS Parametric distribution function	Exponential		Log-logistic		[redacted]
PFS Parametric distribution function	Exponential		Gompertz		[redacted]
PFS data set	Investigator		IRC		[redacted]
PFS treatment effect	9 years	No finite duration	5 years	[redacted]	[redacted]
PPS early PD	GALLIUM	PRIMA			[redacted]
PPS early & late PD pooled	Early/late	PRIMA Pooled	GALLIUM Pooled	[redacted]	[redacted]
Discount rate cost & effect	3.50%		1.5%		[redacted]
Time horizon (years)	40		30		[redacted]

Figure 37: Tornado diagram for base case

[Figure redacted]

5.8.7 For technologies whose final price or acquisition cost has not been confirmed, sensitivity analysis should be done over a plausible range of prices. This may also include the price of a comparator that includes a confidential patient access scheme.

Not applicable.

Scenario analysis

5.8.8 Sensitivity analysis should be used to explore uncertainty around the structural assumptions used in the analysis. Analysis of a representative range of plausible scenarios should be presented and each alternative analysis should present separate results.

Alternative PFS and PPS assumptions: the following key assumptions were different in the de novo model compared to the latest economic analysis of rituximab in combination with chemotherapy in the first-line treatment of FL by Papaioannou or Dewilde: firstly, the Log-normal PFS extrapolation function with no limit on the duration of treatment effect was used in Papaioannou et al. Secondly, PPS was not explicitly dependent on time to progression after first-line treatment. To investigate the impact of these assumptions in the submission model, a Log-normal PFS extrapolation model (with no limit on the duration of treatment effect) was used. Probability of death in early and late PD was assumed to be the same and the pooled early and late post-progression mortality in PRIMA, essentially resulting in a structure with one PD state only. In addition, this scenario included adjusting utilities for age.

Assumptions of equal QALYs and costs post progression: as there is uncertainty around the future QALYs gained and costs post-progression, a simplified scenario, proposed by the ERG for TA251 in chronic myeloid leukaemia (Pavey, 2012), assumes equal QALYs gained and costs between the two arms post progression.

5.8.9 Present the results of scenario analysis. Include details of structural sensitivity analysis.

Alternative PFS and PPS assumptions: although the use of different PFS extrapolation and PPS assumptions in the literature, as described in 5.8.8., leads to an increased gain in life years in PFS (3.75 years median, 2.63 mean undiscounted) compared to the base case (2.75 years median, 1.92 mean undiscounted) the resulting overall life years gained (1.42 mean undiscounted) is similar to the base case (1.45 mean undiscounted) resulting in an ICER comparable to the base case (Table 91).

Table 91: Scenario analysis – alternative PFS and PPS assumptions

Technologies	Total Costs (£)	Total LYG	Total QALYs	Inc Costs (£)	Inc LYG	Inc QALYs	ICER (£)
G-chemo+G	[redacted]	13.66	10.02				
R-chemo+R	[redacted]	12.89	9.30	[redacted]	0.76	0.72	[redacted]

Values in the table are discounted and half cycle corrected

Assumptions of equal QALYs and costs post progression: in this scenario, only the cost difference and QALYs gained in PFS are considered (assuming no difference in cost and QALYs gained post progression) which resulted in a ICER lower than the base-case model (Table 92).

Table 92: Scenario analysis – assumption on equal post progression QALY and cost

Technologies	Total Costs (£)	Total LYG in PFS	Tot QALYs in PFS	Inc Costs (£)	Inc LYG	Inc QALYs	ICER (£)
G-chemo+G	[redacted]	8.77	7.19				
R-chemo+R	[redacted]	7.46	6.12	[redacted]	1.31	1.07	[redacted]

Values in the table are discounted and half cycle corrected

Summary of sensitivity analyses results

5.8.10 Describe the main findings of the sensitivity analyses, highlighting the key drivers of the cost-effectiveness results.

Extensive deterministic sensitivity analyses were conducted by varying individual parameters around using the 10% and 90% percentile from the probabilistic distribution simulation as lower and upper values, respectively. In addition, sensitivity of the results were tested by using alternative sources for utilities, alternative assumptions on time on treatment, administration cost, vial sharing and PFS extrapolation functions. In addition, scenarios investigated the alternative PFS extrapolation functions and alternative PPS assumptions.

The ICERs remained close to the base-case value in most cases. The ICER was most sensitive for the following inputs:

Clinical inputs

The ICER was sensitive to the choice of parametric distribution for PFS. In particular, use of the alternative plausible Log-normal parametric distributions for PFS resulted in lower ICER of £[redacted]/QALY whereas use of the Weibull function increase the ICER to [redacted] QALY. Using the secondary endpoint of IRC assessed PFS resulted in an ICER of [redacted] /QALY. In addition, the ICER was sensitive to the duration of treatment effect. Shortening the parameter to 5 years duration (longest follow up in GALLIUM) of effect resulted in an ICER of [redacted] /QALY. However, a clinical more plausible assumption is that there is no finite duration of the treatment on progression in FL and this resulted in an ICER of [redacted]QALY.

Utilities

The ICER was sensitive to the utility in PFS, in particular to the utility in the maintenance/follow up period, as patients are expected to gain a significant amount of time in PFS on G-chemo+G versus R-chemo+R. However, the base case value for PFS from GALLIUM was derived from a relatively large sample of questionnaires and was therefore

considered robust. The ICER was mainly sensitive to the assumptions on utility in the PD states: using the very conservative values from GALLIUM that did not contain significant follow up beyond progression in the trial, increased the ICER to [redacted] /QALY. In addition, adjusting utilities for an age dependent decline in line with the general UK population that had not been observed in the baseline utilities in GALLIUM increased the ICER to £[redacted]/QALY.

Cost

The ICER was mainly sensitive to the drug acquisition costs. Using time on treatment as per protocol (i.e. assuming all patients in PFS would receive treatment per protocol while in PFS rather than as observed in GALLIUM) increased the ICER to £[redacted] due to the higher technology costs. However, this scenario would require full adherence to protocol in clinical practice and disregard treatment discontinuation due to tolerability or other reasons. In addition, the increase in ICER relies on assumption that the additional treatment received per protocol would have resulted in no additional clinical benefit compared to the observed benefits in GALLIUM.

Discounting

Due to the indolent nature of the disease, a significant amount of health benefits accrue over a longer time period. The ICER was therefore sensitive to the discount rate and using an alternative value of 1.5% (for costs and health effects) decreased the ICER significantly to £[redacted]/QALY.

5.9 Subgroup analysis

5.9.1 Types of subgroups that are not considered relevant are those based solely on the following factors:

- **Individual utilities for health states and patient preference.**
- **Different treatment costs for individuals according to their social characteristics.**
- **Subgroups specified according to the costs of providing treatment in different locations in England (for example, when the costs of facilities available for providing the technology vary according to location).**

5.9.2 Please specify whether analysis of subgroups was carried out and how these subgroups were identified, referring to the scope and decision problem specified for the NICE technology appraisal. When specifying how subgroups were identified,

confirm whether they were identified based on a prior expectation of different clinical or cost effectiveness because of known, biologically plausible mechanisms, social characteristics or other clearly justified factors. Cross refer to the clinical effectiveness section 4.7.

No sub-group analysis was undertaken for the economic analysis. There were no subgroups identified in the scope. The sub-group analysis by GALLIUM trial stratification criteria, e.g. by prognostic FLIPI score and chemotherapy regimen, is discussed in section 4.7 and overall the results of the PFS subgroup analyses are consistent with the primary analysis of PFS in the FL population. Moreover, the GALLIUM study was not powered for significance in the pre-specified sub-groups discussed in section 4.7.

5.9.3 Clearly define the characteristics of patients in the subgroup.

See 5.9.2

5.9.4 Describe how the statistical analysis was carried out.

See 5.9.2

5.9.5 If subgroup analyses were done, please present the results in tables similar to those in section 5.7.

See 5.9.2

5.9.6 Identify any obvious subgroups that were not considered and explain why. Please refer to the subgroups identified in the decision problem in section 3.

See 5.9.2

5.10 Validation

Validation of de novo cost-effectiveness analysis

5.10.1 When describing the methods used to validate and quality assure the model, provide:

- **the rationale for using the chosen methods**
- **references to the results produced and cross-references to the evidence identified in the clinical evidence, measurement and valuation of health effects, and cost and healthcare resource sections.**

The model concept with key clinical inputs, assumptions and clinical outputs was presented to a clinical advisory board of nine UK clinicians to ensure face validity of assumptions and main clinical results.

The Excel version of the model was developed within Roche and an external agency checked the technical validity of the model. This technical validation comprised the following areas:

- Checking whether the statistical parameters (SAS outputs) derived from the trial correspond with the data implemented in the model calculations
- Checking for technical programming or calculation errors (this includes the VBA coding)
- Looking for logical errors or common sense issues related to the model structure, assumptions, data inputs, results and graphical representations.

5.11 Interpretation and conclusions of economic evidence

5.11.1 When interpreting and concluding your economic evidence, consider the following:

- **Are the results from this economic evaluation consistent with the published economic literature? If not, why do the results from this evaluation differ, and why should the results in the submission be given more credence than those in the published literature?**
- **Is the economic evaluation relevant to all groups of patients who could potentially use the technology as identified in the decision problem?**
- **How relevant (generalisable) is the analysis to clinical practice in England?**
- **What are the main strengths and weaknesses of the evaluation? How might these affect the interpretation of the results?**
- **What further analyses could be carried out to enhance the robustness or completeness of the results?**

Conclusion of the economic evidence

The GALLIUM trial demonstrated clinical meaningful and statistical significant improvements, reducing the risk of death or progression by 34% in the primary analysis for G-chemo+G compared to R-chemo+R for previously untreated patients with advanced FL. The de novo economic model predicted that this resulted in a median PFS increase of 2.75 years and mean increase in the time spend free of progression of 1.9 years (undiscounted) for G-chemo+G versus R-chemo+R. This PFS benefit translated in to an (undiscounted) overall survival gain of 1.45 years.

The results of the de novo cost effectiveness analysis of G-chemo+G show that it is both more effective (0.77 QALYs gained) and more costly [redacted]) than R-chemo+R with an ICER of [redacted]QALY.

Relevance to the licensed patient population

The economic evaluation is based on the GALLIUM trial which is representative of the licensed patient population.

Relevance to the UK

All resource use, costs and utility values were taken from sources relevant to England. The NICE 'Single technology appraisal: User guide for company evidence submission template' and the 'Guide to the methods of technology appraisal 2013' were followed throughout. Every step possible was taken to ensure that the analysis undertaken was as pragmatic as possible and accurately estimated the likely costs and health outcomes associated with an average English patient with advanced FL who would currently be treated by R-chemo followed by R maintenance. The main clinical inputs of the model were derived from patient level data from the GALLIUM study that recruited 293 UK patients in 29 centres in the UK. The results produced therefore have strong applicability to an English clinical setting.

Strengths of the economic evaluation

- The economic model is based on the GALLIUM trial, a large, robust and well conducted study in a patient population which is representative of the licensed indication.
- A significant proportion of patients in GALLIUM were from UK centres.
- OS was modelled from PFS and PPS. GALLIUM PFS data was extrapolated using parametric functions. PPS was modelled dependent on time to first progression, consistent with long term follow up data on outcomes on R-chemo and R-chemo+R cohorts in the literature.
- Extensive sensitivity analysis has been performed on the model parameters.
- Although modelling required to account for outcomes and costs of later treatment lines, the modelling approach required relatively few external data sources and sensitivity analyses showed that the conclusions were robust against alternative assumptions.

- Utility values in PFS were available from the GALLIUM study and from a well-conducted UK cross-sectional study.

Areas of weakness or uncertainty

- Due to the indolent nature of FL and the first-line treatment setting, PFS data in GALLIUM was immature and there were few OS events. A Markov approach was required to extrapolate long terms outcomes and costs.
- Trial based EQ-5D utility values were not available significantly beyond progression.

Potential for further analysis

- Long-term follow up data from GALLIUM may reduce some of the uncertainty in the analysis. However, as the case of PRIMA demonstrates, median PFS is unlikely to be reached with less than 10 years follow up.
- Improved measurement of longitudinal utility values post-progression could reduce some uncertainty in the economic analysis.

6. Assessment of factors relevant to the NHS and other parties

6.1 The purpose of this section is to provide an analysis of any factors relevant to the NHS and other parties that may fall outside the remit of the assessments of clinical and cost effectiveness. This will allow subsequent evaluation of the budget impact analysis. Such factors might include issues relating to service organisation and provision, resource allocation and equity, societal or ethical issues, plus any impact on patients or carers. Provide the information specified in sections 6.2–6.10.

6.2 State how many people are eligible for treatment in England. Present results for the full Marketing Authorisation or CE marking and for any subgroups considered. Also present results for the subsequent 5 years.

In England, 2,142 new cases of FL were reported in 2015 (Office for National Statisitics, 2017), with an increase of the incidence of approximately 1% per annum over the last 5 years. However, not all patients are advanced stage or require treatment. According to HMRN data, 47% of patients were treated with chemotherapy regimens after diagnosis (Haematological Malignancy Research Network, 2014). In addition to these patients, patients progressing/developing symptoms after observation ('watch and wait') may also require treatment (Ardeshna et al., 2014). The number of patients starting treatment with R-chemo+R (based on 2015 FL incidence) is estimated at 1152 per year in Table 93.

Table 93: Estimate of the eligeble population in England (2015)

Step	Population	Proportion	No. of People	Source/Assumption
1	Follicular Lymphoma diagnosis in England 2014	100%	2,142	Cancer Statistics 2015 (ONS 2017)
2	Active monitoring (watch and wait) after diagnosis	39% (of 1)	835	HMRN 2014
3	Requiring treatment after watch and wait	54% (of 2)	451	Assumption based on Adershna et al. 2014 (54% requiring treatment after watch and wait during trial follow up).
4	Treated with chemotherapy regimen after diagnosis	47% (of 1)	1007	HMRN 2014
5	Treated with chemotherapy regimen first-line	N/A	1458	Sum of (2) & (3)
6	Treated with MabThera based chemotherapy (induction) 1st line	79% (of 5)	1152	HMRN 2014

Table 94 shows the estimated incidence of previously untreated patients starting treatment with an anti-CD20 in combination with chemotherapy in England based on the estimate in Table 93 with and assuming 1% increase per annum.

Table 94: Eligible population by year in England and Wales

Year	2018	2019	2020	2021	2022
Eligible population	1187	1199	1211	1223	1235

6.3 Explain any assumptions that were made about current treatment options and uptake of technologies.

The analysis considers the difference in budget between G-chemo+G acompared to R-chemo+R. Costs for G-chemo+G and R-chemo+R were based on the outputs of the economic model.

6.4 When relevant, explain any assumptions that were made about market share in England.

The market share estimates are presented in Table 95 ([redacted]

Table 95: Market share assumptions by year

Year	1	2	3	4	5
% people treated with G-benda+G (number starting each year)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

6.5 In addition to technology costs please consider other significant costs associated with treatment that may be of interest to commissioners (for example, administration costs, monitoring costs and the costs of managing adverse reactions).

Drug administration, adverse event, supportive care and subsequent treatment costs were included in the budget impact calculation.

6.6 State what unit costs were assumed and how they were calculated. If unit costs used in health economic modelling were not based on national reference costs or the payment-by-results tariff, explain how a cost for the activity was calculated.

The budget impact calculations are based on the output of the economic model.

6.7 If there were any estimates of resource savings, explain what they were and when they are likely to be made.

Supportive care and subsequent treatment costs were lower for G-chemo+G from year 3 due to fewer patients progressing compared current practice (R-chemo+R).

6.8 State the estimated annual budget impact on the NHS in England.

The annual extimated budget impact for England is shown in the table below.

Table 96: Budget impact by year

Year	1	2	3	4	5
Budget impact - drug cost (£)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Budget impact - non-drug cost (£)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total budget impact (£)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

6.9 Identify any other opportunities for resource savings or redirection of resources that it has not been possible to quantify.

None identified.

6.10 Highlight the main limitations within the budget impact analysis

For budgeting purposes, it was assumed that all patients estimated to start treatment according to Table 95 would start at the beginning of each year.

7. References

ADVANI, R., ROSENBERG, S. A. & HORNING, S. J. 2004. Stage I and II follicular non-Hodgkin's lymphoma: long-term follow-up of no initial therapy. *J Clin Oncol,* 22**,** 1454-9.

AL-TOURAH, A. J., GILL, K. K., CHHANABHAI, M., HOSKINS, P. J., KLASA, R. J., SAVAGE, K. J., SEHN, L. H., SHENKIER, T. N., GASCOYNE, R. D. & CONNORS, J. M. 2008. Population-Based Analysis of Incidence and Outcome of Transformed Non-Hodgkin's Lymphoma. *Journal of Clinical Oncology,* 26**,** 5165-5169.

ALONSO*, S., ALCOCEBA*, M., MAGNANO, L., ANDRADE, M., GARCÍA-ÁLVAREZ, M., MERCADAL, S., RODRÍGUEZ, G., GARCIA, O., SANCHO, J.-M., SALAR, A., PASAROLLS, F., TEROL, M. J., GRANDE, C., GONZÁLEZ DE VILLAMBROSIA, S., CORDOBA, R., NOVELLI, S., LÓPEZ, L., MONTALBÁN, C., DE CABO, E., INFANTE, M. S., PARDAL, E., LOPEZ-JIMENEZ, J., ANTELO, B., ARRANZ, R., GONZÁLEZ, M., MARTÍN, A., LÓPEZ-GUILLERMO‡, A. & CABALLERO‡, M. D. 2015. Incidence, Risk Factors and Prognosis of Transformation in Follicular Lymphoma: a Multicentre Retrospective Analysis of 1763 Patients from the Geltamo Spanish Lymphoma Cooperative Group. *Blood,* 126**,** 3944-3944.

ANDRADE-CAMPOS, M. M., MONTES-LIMÓN, A. E., SORO-ALCUBIERRE, G., GRASA, J. M., LOPEZ-GÓMEZ, L., BARINGO, T. & GIRALDO, P. 2014. Long-term efficacy of 90Y ibritumomab tiuxetan therapy in follicular non-Hodgkin lymphoma and health-related quality of life. *Annals of hematology,* 93**,** 1985-1992.

ARA, R. & BRAZIER, J. E. 2011. Using Health State Utility Values from the General Population to Approximate Baselines in Decision Analytic Models when Condition-Specific Data are Not Available. *Value in Health,* 14**,** 539-545.

ARDESHNA, K. M., QIAN, W., SMITH, P., BRAGANCA, N., LOWRY, L., PATRICK, P., WARDEN, J., STEVENS, L., POCOCK, C. F., MIALL, F., CUNNINGHAM, D., DAVIES, J., JACK, A., STEPHENS, R., WALEWSKI, J., FERHANOGLU, B., BRADSTOCK, K. & LINCH, D. C. 2014. Rituximab versus a watch-and-wait approach in patients with advanced-stage, asymptomatic, non-bulky follicular lymphoma: an open-label randomised phase 3 trial. *Lancet Oncol,* 15**,** 424-35.

BEC M, COGNET M, TAIEB V, PACOU M & GAUTHIER A 2014. French Utility Elicitation in Previously Treated European Patients with Indolent Non-Hodgkin Lymphoma (INHL). *Value Health,* 17**,** A646.

BELLO, C., ZHANG, L. & NAGHASHPOUR, M. 2012. Follicular lymphoma: current management and future directions. *Cancer Control,* 19**,** 187-95.

BEUSTERIEN, K. M., DAVIES, J., LEACH, M., MEIKLEJOHN, D., GRINSPAN, J. L., O'TOOLE, A. & BRAMHAM-JONES, S. 2010. Population preference values for treatment outcomes in chronic lymphocytic leukaemia: a cross-sectional utility study. *Health Qual Life Outcomes,* 8**,** 50.

BLAES A, MA L, ZHANG Y & PETERSON B 2011. Quality of life appears similar between survivors of indolent and aggressive non-Hodgkin lymphoma. *Leuk Lymphoma,* 52**,** 2105-10.

BRITISH NATIONAL FORMULARY. 2017. Available: https://www.medicinescomplete.com/mc/bnf/current/ [Accessed April 2017.

CANCER RESEARCH UK. 2017. *Survival statistics for non Hodgkin lymphoma* [Online]. Available: http://www.cancerresearchuk.org/about-cancer/type/non-hodgkins-lymphoma/treatment/statistics-and-outlook-for-non-hodgkins-lymphoma April 2017].

CASULO, C., BURACK, W. R. & FRIEDBERG, J. W. 2015a. Transformed follicular non-Hodgkin lymphoma. *Blood,* 125**,** 40-7.

CASULO, C., BYRTEK, M., DAWSON, K. L., ZHOU, X., FARBER, C. M., FLOWERS, C. R., HAINSWORTH, J. D., MAURER, M. J., CERHAN, J. R., LINK, B. K., ZELENETZ, A. D. & FRIEDBERG, J. W. 2015b. Early Relapse of Follicular Lymphoma After Rituximab Plus Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone Defines Patients at High Risk for Death: An Analysis From the National LymphoCare Study. *J Clin Oncol,* 33**,** 2516-22.

CENTRE FOR REVIEWS AND DISSEMINATION. *CRD's guidance for undertaking reviews in health care* [Online]. Available: https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf April 2017].

CHESON BD, PFISTNER B, JUWEID ME, GASCOYNE, R., SPECHT L, HORNING SJ, COIFFIER B, FISHER RI, HAGENBEEK A, ZUCCA E, ROSEN ST, STROOBANTS S, LISTER TA, HOPPE RT, DREYLING M, TOBINAI K, VOSE JM, CONNORS JM, FEDERICO M & DIEHL V 2007. Revised Response Criteria for Malignant Lymphoma. *Journal of Clinical Oncology,* 25**,** 579-586.

CHEUNG, M. C., IMRIE, K. R., FRIEDLICH, J., BUCKSTEIN, R., LATHIA, N. & MITTMANN, N. 2009. The impact of follicular (FL) and other indolent non-Hodgkin's lymphomas (NHL) on work productivity-a preliminary analysis. *Psychooncology,* 18**,** 554-9.

CLINICAL TRIALS.GOV. *A Study of Obinutuzumab (RO5072759) Plus Chemotherapy in Comparison With MabThera/Rituxan (Rituximab) Plus Chemotherapy Followed by GA101 or MabThera/Rituxan Maintenance in Patients With Untreated Advanced Indolent Non-Hodgkin's Lymphoma (GALLIUM)* [Online]. Available: https://clinicaltrials.gov/ct2/show/NCT01332968?term=21223&rank=1 Accessed January 2017].

CURTIS, L. B., A. 2016. Unit Costs of Health and Social Care 2016. Canterbury: Personal Social Services Research Unit, University of Kent.

CZUCZMAN, M. S., FAYAD, L., DELWAIL, V., CARTRON, G., JACOBSEN, E., KULICZKOWSKI, K., LINK, B. K., PINTER-BROWN, L., RADFORD, J., HELLMANN, A., GALLOP-EVANS, E., DIRIENZO, C. G., GOLDSTEIN, N., GUPTA, I., JEWELL, R. C., LIN, T. S., LISBY, S., SCHULTZ, M., RUSSELL, C. A., HAGENBEEK, A. & STUDY, I. 2012. Ofatumumab monotherapy in rituximab-refractory follicular lymphoma: results from a multicenter study. *Blood,* 119**,** 3698-704.

DALLE, S., RESLAN, L., BESSEYRE DE HORTS, T., HERVEAU, S., HERTING, F., PLESA, A., FRIESS, T., UMANA, P., KLEIN, C. & DUMONTET, C. 2011. Preclinical studies on the mechanism of action and the anti-lymphoma activity of the novel anti-CD20 antibody GA101. *Mol Cancer Ther,* 10**,** 178-85.

DEPARTMENT OF HEALTH. 2016. *Drugs and pharmaceutical electronic market information (eMit)* [Online]. Commercial Medicines Unit (CMU). Available: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit [Accessed March 2017 2017].

DEPARTMENT OF HEALTH. 2017. *NHS Reference Costs 2015-16* [Online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/577083/Reference_Costs_2015-16.pdf [Accessed April 2017.

DEWILDE, S., WOODS, B., CASTAIGNE, J. G., PARKER, C. & DUNLOP, W. 2014. Bendamustine-rituximab: a cost-utility analysis in first-line treatment of indolent non-Hodgkin's lymphoma in England and Wales. *J Med Econ,* 17**,** 111-24.

DREYLING, M. 2009. Newly diagnosed and relapsed follicular lymphoma: ESMO clinical recommendations for diagnosis, treatment and follow-up. *Ann Oncol,* 20 Suppl 4**,** 119-20.

DREYLING, M., GHIELMINI, M., RULE, S., SALLES, G., VITOLO, U. & LADETTO, M. 2016. Newly diagnosed and relapsed follicular lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol,* 27**,** v83-v90.

DUNDAR, Y., BAGUST, A., HOUNSOME, J., MCLEOD, C., BOLAND, A., DAVIS, H., WALLEY, T. & DICKSON, R. 2009. Rituximab for the first-line treatment of stage III/IV follicular non-Hodgkin's lymphoma. *Health technology assessment (Winchester, England),* 13 Suppl 1**,** 23-8.

EUROPEAN MEDICINES AGENCY 2016. European Public Assessment Report - Obinutuzumab.

F. HOFFMANN-LA ROCHE LTD 2016. Primary Clinical Study Report - BO21223, Report Number 1067980.

F. HOFFMANN-LA ROCHE LTD 2017. Obinutuzumab Summary of Product Characteristics [DRAFT].

FAHAM, M., ZHENG, J., MOORHEAD, M., CARLTON, V. E., STOW, P., COUSTAN-SMITH, E., PUI, C. H. & CAMPANA, D. 2012. Deep-sequencing approach for minimal residual disease detection in acute lymphoblastic leukemia. *Blood,* 120**,** 5173-80.

FEDERICO, M., BELLEI, M., MARCHESELLI, L., LUMINARI, S., LOPEZ-GUILLERMO, A., VITOLO, U., PRO, B., PILERI, S., PULSONI, A., SOUBEYRAN, P., CORTELAZZO, S., MARTINELLI, G., MARTELLI, M., RIGACCI, L., ARCAINI, L., DI RAIMONDO, F., MERLI, F., SABATTINI, E., MCLAUGHLIN, P. & SOLAL-CELIGNY, P. 2009. Follicular lymphoma international prognostic index 2: a new prognostic index for follicular lymphoma developed by the international follicular lymphoma prognostic factor project. *J Clin Oncol,* 27**,** 4555-62.

FISCHBACH, W., SCHRAMM, S. & GOEBELER, E. 2011. Outcome and quality of life favour a conservative treatment of patients with primary gastric lymphoma. *Z Gastroenterol,* 49**,** 430-5.

FISHER, R. I., LEBLANC, M., PRESS, O. W., MALONEY, D. G., UNGER, J. M. & MILLER, T. P. 2005. New treatment options have changed the survival of patients with follicular lymphoma. *J Clin Oncol,* 23**,** 8447-52.

FREEDMAN, A. 2015. Follicular lymphoma: 2015 update on diagnosis and management. *Am J Hematol,* 90**,** 1171-8.

FRIEDBERG, J. W., BYRTEK, M., LINK, B. K., FLOWERS, C., TAYLOR, M., HAINSWORTH, J., CERHAN, J. R., ZELENETZ, A. D., HIRATA, J. & MILLER, T. P. 2012. Effectiveness of first-line management strategies for stage I follicular lymphoma: analysis of the National LymphoCare Study. *J Clin Oncol,* 30**,** 3368-75.

FRIEDBERG, J. W., COHEN, P., CHEN, L., ROBINSON, K. S., FORERO-TORRES, A., LA CASCE, A. S., FAYAD, L. E., BESSUDO, A., CAMACHO, E. S., WILLIAMS, M. E., VAN DER JAGT, R. H., OLIVER, J. W. & CHESON, B. D. 2008. Bendamustine in patients with rituximab-refractory indolent and transformed non-Hodgkin's lymphoma: results from a phase II multicenter, single-agent study. *J Clin Oncol,* 26**,** 204-10.

FRIEDLICH, J. D., CHEUNG, M. C., IMRIE, K. R., HALES, B., MITTMANN, N. & BUCKSTEIN, R. 2006. Discrimination of Health States in Follicular Lymphoma with Utilities Derived from the EuroQOL EQ5D Instrument. *Blood,* 108**,** 3329-3329.

GODON, A., MOREAU, A., TALMANT, P., BARANGER-PAPOT, L., GENEVIEVE, F., MILPIED, N., ZANDECKI, M. & AVET-LOISEAU, H. 2003. Is t(14;18)(q32;q21) a constant finding in follicular lymphoma? An interphase FISH study on 63 patients. *Leukemia,* 17**,** 255-9.

GOEDE, V., FISCHER, K., BUSCH, R., ENGELKE, A., EICHHORST, B., WENDTNER, C. M., CHAGOROVA, T., DE LA SERNA, J., DILHUYDY, M. S., ILLMER, T., OPAT, S., OWEN, C. J., SAMOYLOVA, O., KREUZER, K. A., STILGENBAUER, S., DOHNER, H., LANGERAK, A. W., RITGEN, M., KNEBA, M., ASIKANIUS, E., HUMPHREY, K., WENGER, M. & HALLEK, M. 2014. Obinutuzumab plus chlorambucil in patients with CLL and coexisting conditions. *N Engl J Med,* 370**,** 1101-10.

GOEDE, V., KLEIN, C. & STILGENBAUER, S. 2015. Obinutuzumab (GA101) for the treatment of chronic lymphocytic leukemia and other B-cell non-hodgkin's lymphomas: a glycoengineered type II CD20 antibody. *Oncol Res Treat,* 38**,** 185-92.

GOLAY, J., DA ROIT, F., BOLOGNA, L., FERRARA, C., LEUSEN, J. H., RAMBALDI, A., KLEIN, C. & INTRONA, M. 2013. Glycoengineered CD20 antibody obinutuzumab activates neutrophils and mediates phagocytosis through CD16B more efficiently than rituximab. *Blood,* 122**,** 3482-91.

GREENHALGH, J., BAGUST, A., BOLAND, A., BLUNDELL, M., OYEE, J., BEALE, S., DUNDAR, Y., HOCKENHULL, J., PROUDLOVE, C. & CHU, P. 2013. Rituximab for the first-line maintenance treatment of follicular non-Hodgkin's lymphoma : a NICE single technology appraisal. *Pharmacoeconomics,* 31**,** 403-13.

GRIGG, A., DYER, M. J. S., GONZÁLEZ DÍAZ, M., DREYLING, M., RULE, S., LEI, G., KNAPP, A., WASSNER-FRITSCH, E. & MARLTON, P. 2016. Safety and efficacy of Obinutuzumab with CHOP or bendamustine in previously untreated follicular lymphoma. *Haematologica*.

HAEMATOLOGICAL MALIGNANCY RESEARCH NETWORK 2014. Patient's age and treatment for haematological malignancy: a report from the Haematological Malignancy Research Network (HMRN).

HAEMATOLOGICAL MALIGNANCY RESEARCH NETWORK. 2017a. *FL Incidence Statistics* [Online]. Available: https://www.hmrn.org/statistics/incidence April 2017].

HAEMATOLOGICAL MALIGNANCY RESEARCH NETWORK. 2017c. *FL Prevalence Statistics* [Online]. Available: https://www.hmrn.org/statistics/prevalence April 2017].

HAEMATOLOGICAL MALIGNANCY RESEARCH NETWORK. 2017d. *FL Survival Statistics* [Online]. Available: https://www.hmrn.org/statistics/survival April 2017].

HEROLD, M., HAAS, A., SROCK, S., NESER, S., AL-ALI, K. H., NEUBAUER, A., DOLKEN, G., NAUMANN, R., KNAUF, W., FREUND, M., ROHRBERG, R., HOFFKEN, K., FRANKE, A., ITTEL, T., KETTNER, E., HAAK, U., MEY, U., KLINKENSTEIN, C., ASSMANN, M. & VON GRUNHAGEN, U. 2007. Rituximab added to first-line mitoxantrone, chlorambucil, and prednisolone chemotherapy followed by interferon maintenance prolongs survival in patients with advanced follicular lymphoma: an East German Study Group Hematology and Oncology Study. *J Clin Oncol,* 25**,** 1986-92.

HIDDEMANN, W. & CHESON, B. D. 2014. How we manage follicular lymphoma. *Leukemia,* 28**,** 1388-95.

HIDDEMANN, W., KNEBA, M., DREYLING, M., SCHMITZ, N., LENGFELDER, E., SCHMITS, R., REISER, M., METZNER, B., HARDER, H., HEGEWISCH-BECKER, S., FISCHER, T., KROPFF, M., REIS, H. E., FREUND, M., WORMANN, B., FUCHS, R., PLANKER, M., SCHIMKE, J., EIMERMACHER, H., TRUMPER, L., ALDAOUD, A., PARWARESCH, R. & UNTERHALT, M. 2005. Frontline therapy with rituximab added to the combination of cyclophosphamide, doxorubicin, vincristine, and prednisone (CHOP) significantly improves the outcome for patients with advanced-stage follicular lymphoma compared with therapy with CHOP alone: results of a prospective randomized study of the German Low-Grade Lymphoma Study Group. *Blood,* 106**,** 3725-32.

HOCHSTER, H., WELLER, E., GASCOYNE, R. D., HABERMANN, T. M., GORDON, L. I., RYAN, T., ZHANG, L., COLOCCI, N., FRANKEL, S. & HORNING, S. J. 2009. Maintenance rituximab after cyclophosphamide, vincristine, and prednisone prolongs progression-free survival in advanced indolent lymphoma: results of the randomized phase III ECOG1496 Study. *J Clin Oncol,* 27**,** 1607-14.

HORNING, S. J., YOUNES, A., JAIN, V., KROLL, S., LUCAS, J., PODOLOFF, D. & GORIS, M. 2005. Efficacy and safety of tositumomab and iodine-131 tositumomab (Bexxar) in B-cell lymphoma, progressive after rituximab. *J Clin Oncol,* 23**,** 712-9.

JUWEID, M. E., STROOBANTS, S., HOEKSTRA, O. S., MOTTAGHY, F. M., DIETLEIN, M., GUERMAZI, A., WISEMAN, G. A., KOSTAKOGLU, L., SCHEIDHAUER, K., BUCK, A., NAUMANN,

R., SPAEPEN, K., HICKS, R. J., WEBER, W. A., RESKE, S. N., SCHWAIGER, M., SCHWARTZ, L. H., ZIJLSTRA, J. M., SIEGEL, B. A. & CHESON, B. D. 2007. Use of positron emission tomography for response assessment of lymphoma: consensus of the Imaging Subcommittee of International Harmonization Project in Lymphoma. *J Clin Oncol,* 25**,** 571-8.

KAHL, B. S., BARTLETT, N. L., LEONARD, J. P., CHEN, L., GANJOO, K., WILLIAMS, M. E., CZUCZMAN, M. S., ROBINSON, K. S., JOYCE, R., VAN DER JAGT, R. H. & CHESON, B. D. 2010. Bendamustine is effective therapy in patients with rituximab-refractory, indolent B-cell non-Hodgkin lymphoma: results from a Multicenter Study. *Cancer,* 116**,** 106-14.

KORSZUN, A., SARKER, S. J., CHOWDHURY, K., CLARK, C., GREAVES, P., JOHNSON, R., KINGSTON, J., LEVITT, G., MATTHEWS, J., WHITE, P., LISTER, A. & GRIBBEN, J. 2014. Psychosocial factors associated with impact of cancer in longterm haematological cancer survivors. *Br J Haematol,* 164**,** 790-803.

LADETTO, M., LOBETTI-BODONI, C., MANTOAN, B., CECCARELLI, M., BOCCOMINI, C., GENUARDI, E., CHIAPPELLA, A., BALDINI, L., ROSSI, G., PULSONI, A., DI RAIMONDO, F., RIGACCI, L., PINTO, A., GALIMBERTI, S., BARI, A., ROTA-SCALABRINI, D., FERRARI, A., ZAJA, F., GALLAMINI, A., SPECCHIA, G., MUSTO, P., ROSSI, F. G., GAMBA, E., EVANGELISTA, A., VITOLO, U. & FONDAZIONE ITALIANA, L. 2013. Persistence of minimal residual disease in bone marrow predicts outcome in follicular lymphomas treated with a rituximab-intensive program. *Blood,* 122**,** 3759-66.

LATIMER, N. R. 2013. Survival analysis for economic evaluations alongside clinical trials--extrapolation with patient-level data: inconsistencies, limitations, and a practical guide. *Med Decis Making,* 33**,** 743-54.

LEVY, V., PORCHER, R., LEBLOND, V., FERMAND, J., CAZIN, B., MALOISEL, F., HAROUSSEAU, J., REMENIERAS, L., GUIBON, O. & CHEVRET, S. 2001. Evaluating treatment strategies in advanced Waldenström macroglobulinemia: use of quality-adjusted survival analysis. *Leukemia,* 15**,** 1466.

LEWIS G, MARCUS R, PROCTOR S. J, GYLDMARK M, CREEDEN J, AULTMAN R & JOST F 2006. The Cost-Effectiveness of Rituximab, Cyclophosphamide, Vincristine and Prednisolone (RCVP) Compared with CVP for the Treatment of Follicular Non-Hodgkin's Lymphoma (NHL) in the UK. *Blood,* 108**,** 345.

LINK, B. K., MAURER, M. J., NOWAKOWSKI, G. S., ANSELL, S. M., MACON, W. R., SYRBU, S. I., SLAGER, S. L., THOMPSON, C. A., INWARDS, D. J., JOHNSTON, P. B., COLGAN, J. P., WITZIG, T. E., HABERMANN, T. M. & CERHAN, J. R. 2013. Rates and outcomes of follicular lymphoma transformation in the immunochemotherapy era: a report from the University of Iowa/MayoClinic Specialized Program of Research Excellence Molecular Epidemiology Resource. *J Clin Oncol,* 31**,** 3272-8.

LOBETTI-BODONI, C., MANTOAN, B., MONITILLO, L., GENUARDI, E., DRANDI, D., BARBERO, D., BERNOCCO, E., BOCCADORO, M. & LADETTO, M. 2013. Clinical implications and prognostic role of minimal residual disease detection in follicular lymphoma. *Ther Adv Hematol,* 4**,** 189-98.

MA, S. 2012. Risk Factors of Follicular Lymphoma. *Expert Opin Med Diagn,* 6**,** 323-333.

MADSEN C, ROOST CLAUSEN M, LUDVIGSEN M, PLESNAR TL & F, D. A. Upfront Rituximab Maintenance after Induction Therapy Improves Outcome and Reduces the Risk of Histological Transformation in Patients with Follicular Lymphoma – Real World Data from a Danish Population-Based Cohort. American Society of Hematology, 2016.

MARCUS R, DAVIES A, ANDO K, KLAPPER W, OPAT S, OWEN C, PHILLIPS E, SANGHA R, SCHLAG R, SEYMOUR JF, TOWNSEND W, TRNĚNÝ M, WENGER M, FINGERLE-ROWSON G, RUFIBACH K, MOORE T, HEROLD M & HIDDEMANN W. Obinutuzumab-based induction and maintenance prolongs progression-free survival (PFS) in patients with previously untreated follicular

lymphoma: primary results of the randomized Phase III GALLIUM study. American Society of Hematology, 2016.

MARCUS, R., IMRIE, K., BELCH, A., CUNNINGHAM, D., FLORES, E., CATALANO, J., SOLAL-CELIGNY, P., OFFNER, F., WALEWSKI, J., RAPOSO, J., JACK, A. & SMITH, P. 2005. CVP chemotherapy plus rituximab compared with CVP as first-line treatment for advanced follicular lymphoma. *Blood,* 105**,** 1417-23.

MARCUS, R., IMRIE, K., SOLAL-CELIGNY, P., CATALANO, J. V., DMOSZYNSKA, A., RAPOSO, J. C., OFFNER, F. C., GOMEZ-CODINA, J., BELCH, A., CUNNINGHAM, D., WASSNER-FRITSCH, E. & STEIN, G. 2008. Phase III study of R-CVP compared with cyclophosphamide, vincristine, and prednisone alone in patients with previously untreated advanced follicular lymphoma. *J Clin Oncol,* 26**,** 4579-86.

MAURER, M. J., BACHY, E., GHESQUIERES, H., ANSELL, S. M., NOWAKOWSKI, G. S., THOMPSON, C. A., INWARDS, D. J., ALLMER, C., CHASSAGNE-CLEMENT, C., NICOLAS-VIRELIZIER, E., SEBBAN, C., LEBRAS, L., SARKOZY, C., MACON, W. R., FELDMAN, A. L., SYRBU, S. I., TRAVERSE-GLEHAN, A., COIFFIER, B., SLAGER, S. L., WEINER, G. J., WITZIG, T. E., HABERMANN, T. M., SALLES, G., CERHAN, J. R. & LINK, B. K. 2016. Early event status informs subsequent outcome in newly diagnosed follicular lymphoma. *Am J Hematol,* 91**,** 1096-1101.

MCNAMARA, C., DAVIES, J., DYER, M., HOSKIN, P., ILLIDGE, T., LYTTELTON, M., MARCUS, R., MONTOTO, S., RAMSAY, A., WONG, W. L. & ARDESHNA, K. 2012. Guidelines on the investigation and management of follicular lymphoma. *Br J Haematol,* 156**,** 446-67.

MEDSCAPE. 2016. *Follicular Lymphoma Management Overview* [Online]. Available: http://emedicine.medscape.com/article/203268-overview April 2017].

MOSSNER, E., BRUNKER, P., MOSER, S., PUNTENER, U., SCHMIDT, C., HERTER, S., GRAU, R., GERDES, C., NOPORA, A., VAN PUIJENBROEK, E., FERRARA, C., SONDERMANN, P., JAGER, C., STREIN, P., FERTIG, G., FRIESS, T., SCHULL, C., BAUER, S., DAL PORTO, J., DEL NAGRO, C., DABBAGH, K., DYER, M. J., POPPEMA, S., KLEIN, C. & UMANA, P. 2010. Increasing the efficacy of CD20 antibody therapy through the engineering of a new type II anti-CD20 antibody with enhanced direct and immune effector cell-mediated B-cell cytotoxicity. *Blood,* 115**,** 4393-402.

MURAKAMI, S., KATO, H., HIGUCHI, Y., YAMAMOTO, K., YAMAMOTO, H., SAITO, T., TAJI, H., YATABE, Y., NAKAMURA, S. & KINOSHITA, T. 2016. Prediction of high risk for death in patients with follicular lymphoma receiving rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisolone in first-line chemotherapy. *Ann Hematol,* 95**,** 1259-69.

NAFEES, B., STAFFORD, M., GAVRIEL, S., BHALLA, S. & WATKINS, J. 2008. Health state utilities for non small cell lung cancer. *Health Qual Life Outcomes,* 6**,** 84.

NASTOUPIL, L. J., SINHA, R., BYRTEK, M., ZIEMIECKI, R., TAYLOR, M., FRIEDBERG, J. W., KOFF, J. L., LINK, B. K., CERHAN, J. R., DAWSON, K. L. & FLOWERS, C. R. 2015. Comparison of the effectiveness of frontline chemoimmunotherapy regimens for follicular lymphoma used in the United States. *Leuk Lymphoma,* 56**,** 1295-302.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE 2011. Rituximab for the first-line maintenance treatment of follicular non-Hodgkin's lymphoma [TA226].

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE 2012. Rituximab for the first-line treatment of stage III-IV follicular lymphoma [TA243].

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE 2016a. Haematological cancers: improving outcomes [NG47].

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE 2016d. Non-Hodgkin’s lymphoma: diagnosis and management [NG52].

OFFICE FOR NATIONAL STATISITICS. 2017. *Cancer Registration Statistics, England* [Online]. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/cancerregistrationstatisticscancerregistrationstatisticsengland April 2017].

OLIN, R. L., KANETSKY, P. A., TEN HAVE, T. R., NASTA, S. D., SCHUSTER, S. J. & ANDREADIS, C. 2010. Determinants of the optimal first-line therapy for follicular lymphoma: a decision analysis. *Am J Hematol,* 85**,** 255-60.

PAPAIOANNOU, D., RAFIA, R., RATHBONE, J., STEVENSON, M., BUCKLEY WOODS, H. & STEVENS, J. 2012. Rituximab for the first-line treatment of stage III-IV follicular lymphoma (review of Technology Appraisal No. 110): a systematic review and economic evaluation. *Health Technol Assess,* 16**,** 1-253, iii-iv.

PAVEY, T. H., M.; CIANI, O.; CRATHORNE, L.; JONES-HUGHES, T. 2012. Dasatinib, nilotinib and standard-dose imatinib for the first-line treatment of chronic myeloid leukaemia: systematic reviews and economic analyses. *Health Technology Assessment,* 16.

PETTENGELL, R., DONATTI, C., HOSKIN, P., POYNTON, C., KETTLE, P. J., HANCOCK, B., JOHNSON, S., DYER, M. J., RULE, S., WALKER, M. & WILD, D. 2008. The impact of follicular lymphoma on health-related quality of life. *Ann Oncol,* 19**,** 570-6.

POCOCK, S. J., CLAYTON, T. C. & ALTMAN, D. G. 2002. Survival plots of time-to-event outcomes in clinical trials: good practice and pitfalls. *Lancet,* 359**,** 1686-9.

POTT C, HOSTER E, KEHDEN B, UNTERHALT M, HEROLD M, VAN DER JAGT, R. H., JANSSENS A, KNEBA M, MAYER J, POCOCK C, DANESI N, FINGERLE-ROWSON G, HARBRON C, MUNDT K, MARCUS R & W, H. Minimal residual disease in patients with follicular lymphoma treated as first-line with obinutuzumab or rituximab-based immunochemotherapy in the Phase III GALLIUM study. American Society of Hematology, 2016.

POTT, C., BELADA, D., DANESI, N., FINGERLE-ROWSON, G., GRIBBEN, J., HABRON, C., HOSTER, E., KAHL, B. S., MUNDT, K., SEBBAN, C., SEHN, L. H. & CHESON, B. D. Analysis of minimal residual disease in follicular lymphoma patients in GADOLIN, a Phase III study of obinutuzumab plus bendamustine versus bendamustinein relapsed/refractory indolent non-Hodgkin lymphoma. ASH, 5–8 December 2015 Orlando, Florida.

RAMBALDI, A., LAZZARI, M., MANZONI, C., CARLOTTI, E., ARCAINI, L., BACCARANI, M., BARBUI, T., BERNASCONI, C., DASTOLI, G., FUGA, G., GAMBA, E., GARGANTINI, L., GATTEI, V., LAURIA, F., LAZZARINO, M., MANDELLI, F., MORRA, E., PULSONI, A., RIBERSANI, M., ROSSI-FERRINI, P. L., RUPOLO, M., TURA, S., ZAGONEL, V., ZAJA, F., ZINZANI, P., REATO, G. & FOA, R. 2002. Monitoring of minimal residual disease after CHOP and rituximab in previously untreated patients with follicular lymphoma. *Blood,* 99**,** 856-62.

RANCEA, M., WILL, A., BORCHMANN, P., MONSEF, I., ENGERT, A. & SKOETZ, N. 2014. Sixteenth biannual report of the Cochrane Haematological Malignancies Group: focus on Non-Hodgkin's lymphoma. *J Natl Cancer Inst,* 106.

RAY JA, CARR E, LEWIS G & MARCUS R 2010. An evaluation of the cost-effectiveness of rituximab in combination with chemotherapy for the first-line treatment of follicular non-Hodgkin's lymphoma in the UK. *Value Health,* 13**,** 346-357.

ROCHE PRODUCTS LTD. Data on file [RXUKMABO00894].

SALLES, G. iNHL Satellite Symposium. 21st Congress of the Haematology Association, June 9-12, 2016, 2016 Copenhagen, Denmark.

SALLES, G., MOUNIER, N., DE GUIBERT, S., MORSCHHAUSER, F., DOYEN, C., ROSSI, J. F., HAIOUN, C., BRICE, P., MAHE, B., BOUABDALLAH, R., AUDHUY, B., FERME, C., DARTIGEAS, C., FEUGIER, P., SEBBAN, C., XERRI, L. & FOUSSARD, C. 2008. Rituximab combined with

chemotherapy and interferon in follicular lymphoma patients: results of the GELA-GOELAMS FL2000 study. *Blood,* 112**,** 4824-31.

SALLES, G., SEYMOUR, J. F., OFFNER, F., LOPEZ-GUILLERMO, A., BELADA, D., XERRI, L., FEUGIER, P., BOUABDALLAH, R., CATALANO, J. V., BRICE, P., CABALLERO, D., HAIOUN, C., PEDERSEN, L. M., DELMER, A., SIMPSON, D., LEPPA, S., SOUBEYRAN, P., HAGENBEEK, A., CASASNOVAS, O., INTRAGUMTORNCHAI, T., FERME, C., DA SILVA, M. G., SEBBAN, C., LISTER, A., ESTELL, J. A., MILONE, G., SONET, A., MENDILA, M., COIFFIER, B. & TILLY, H. 2011. Rituximab maintenance for 2 years in patients with high tumour burden follicular lymphoma responding to rituximab plus chemotherapy (PRIMA): a phase 3, randomised controlled trial. *Lancet,* 377**,** 42-51.

SARKOZY, C., TRNENY, M., XERRI, L., WICKHAM, N., FEUGIER, P., LEPPA, S., BRICE, P., SOUBEYRAN, P., GOMES DA SILVA, M., MOUNIER, C., OFFNER, F., DUPUIS, J., CABALLERO, D., CANIONI, D., PAULA, M., DELARUE, R., ZACHEE, P., SEYMOUR, J., SALLES, G. & TILLY, H. 2016. Risk Factors and Outcomes for Patients With Follicular Lymphoma Who Had Histologic Transformation After Response to First-Line Immunochemotherapy in the PRIMA Trial. *J Clin Oncol,* 34**,** 2575-82.

SCHATZ, J. H., ORICCHIO, E., PUVVADA, S. D. & WENDEL, H. G. 2013. Progress against follicular lymphoma. *Curr Opin Hematol,* 20**,** 320-6.

SEHN, L. H., CHUA, N., MAYER, J., DUECK, G., TRNENY, M., BOUABDALLAH, K., FOWLER, N., DELWAIL, V., PRESS, O., SALLES, G., GRIBBEN, J., LENNARD, A., LUGTENBURG, P. J., DIMIER, N., WASSNER-FRITSCH, E., FINGERLE-ROWSON, G. & CHESON, B. D. 2016. Obinutuzumab plus bendamustine versus bendamustine monotherapy in patients with rituximab-refractory indolent non-Hodgkin lymphoma (GADOLIN): a randomised, controlled, open-label, multicentre, phase 3 trial. *Lancet Oncol,* 17**,** 1081-93.

SEHN, L. H., GOY, A., OFFNER, F. C., MARTINELLI, G., CABALLERO, M. D., GADEBERG, O., BAETZ, T., ZELENETZ, A. D., GAIDANO, G., FAYAD, L. E., BUCKSTEIN, R., FRIEDBERG, J. W., CRUMP, M., JAKSIC, B., ZINZANI, P. L., PADMANABHAN IYER, S., SAHIN, D., CHAI, A., FINGERLE-ROWSON, G. & PRESS, O. W. 2015. Randomized Phase II Trial Comparing Obinutuzumab (GA101) With Rituximab in Patients With Relapsed CD20+ Indolent B-Cell Non-Hodgkin Lymphoma: Final Analysis of the GAUSS Study. *J Clin Oncol,* 33**,** 3467-74.

SEYMOUR JF, FEUGIER P, OFFNER F, LOPEZ-GUILLERMO A, BELADA D, XERRI L, BOUABDALLAH R, CATALANO J, PAULINE B, CABALLERO D, HAIOUN C, PEDERSEN LM, DELMER A, SIMPSON D, LEPPA S, SOUBEYRAN P, HAGENBEEK A, CASASNOVAS O, INTRAGUMTORNCHAI T, FERME C, GOMES DA SILVA M, SEBBAN C, LISTER A, JA, E., MILONE G, SONET A, COIFFIER B & TILLY H 2013. Updated 6 Year Follow-Up Of The PRIMA Study Confirms The Benefit Of 2-Year Rituximab Maintenance In Follicular Lymphoma Patients Responding To Frontline Immunochemotherapy. *Blood,* 122**,** 509-509.

SOLAL-CELIGNY, P., ROY, P., COLOMBAT, P., WHITE, J., ARMITAGE, J. O., ARRANZ-SAEZ, R., AU, W. Y., BELLEI, M., BRICE, P., CABALLERO, D., COIFFIER, B., CONDE-GARCIA, E., DOYEN, C., FEDERICO, M., FISHER, R. I., GARCIA-CONDE, J. F., GUGLIELMI, C., HAGENBEEK, A., HAIOUN, C., LEBLANC, M., LISTER, A. T., LOPEZ-GUILLERMO, A., MCLAUGHLIN, P., MILPIED, N., MOREL, P., MOUNIER, N., PROCTOR, S. J., ROHATINER, A., SMITH, P., SOUBEYRAN, P., TILLY, H., VITOLO, U., ZINZANI, P. L., ZUCCA, E. & MONTSERRAT, E. 2004. Follicular lymphoma international prognostic index. *Blood,* 104**,** 1258-65.

SOLIMANDO, A. G., RIBATTI, D., VACCA, A. & EINSELE, H. 2016. Targeting B-cell non Hodgkin lymphoma: New and old tricks. *Leuk Res,* 42**,** 93-104.

SWINBURN, P., LLOYD, A., NATHAN, P., CHOUEIRI, T. K., CELLA, D. & NEARY, M. P. 2010. Elicitation of health state utilities in metastatic renal cell carcinoma. *Curr Med Res Opin,* 26**,** 1091-6.

TAN, D., HORNING, S. J., HOPPE, R. T., LEVY, R., ROSENBERG, S. A., SIGAL, B. M., WARNKE, R. A., NATKUNAM, Y., HAN, S. S., YUEN, A., PLEVRITIS, S. K. & ADVANI, R. H. 2013.

Improvements in observed and relative survival in follicular grade 1-2 lymphoma during 4 decades: the Stanford University experience. *Blood,* 122**,** 981-7.

TARELLA, C., GUELI, A., DELAINI, F., ROSSI, A., BARBUI, A. M., GRITTI, G., BOSCHINI, C., CARACCIOLO, D., BRUNA, R., RUELLA, M., GOTTARDI, D., PASSERA, R. & RAMBALDI, A. 2014. Rate of primary refractory disease in B and T-cell non-Hodgkin's lymphoma: correlation with long-term survival. *PLoS One,* 9**,** e106745.

TOLLEY, K., GOAD, C., YI, Y., MAROUDAS, P., HAIDERALI, A. & THOMPSON, G. 2013. Utility elicitation study in the UK general public for late-stage chronic lymphocytic leukaemia. *Eur J Health Econ,* 14**,** 749-59.

VAN OERS, M. H., KLASA, R., MARCUS, R. E., WOLF, M., KIMBY, E., GASCOYNE, R. D., JACK, A., VAN'T VEER, M., VRANOVSKY, A., HOLTE, H., VAN GLABBEKE, M., TEODOROVIC, I., ROZEWICZ, C. & HAGENBEEK, A. 2006. Rituximab maintenance improves clinical outcome of relapsed/resistant follicular non-Hodgkin lymphoma in patients both with and without rituximab during induction: results of a prospective randomized phase 3 intergroup trial. *Blood,* 108**,** 3295-301.

VAN OERS, M. H., VAN GLABBEKE, M., GIURGEA, L., KLASA, R., MARCUS, R. E., WOLF, M., KIMBY, E., VAN T VEER, M., VRANOVSKY, A., HOLTE, H. & HAGENBEEK, A. 2010. Rituximab maintenance treatment of relapsed/resistant follicular non-Hodgkin's lymphoma: long-term outcome of the EORTC 20981 phase III randomized intergroup study. *J Clin Oncol,* 28**,** 2853-8.

VIDAL, L., GAFTER-GVILI, A., SALLES, G., DREYLING, M. H., GHIELMINI, M., HSU SCHMITZ, S. F., PETTENGELL, R., WITZENS-HARIG, M. & SHPILBERG, O. 2011. Rituximab maintenance for the treatment of patients with follicular lymphoma: an updated systematic review and meta-analysis of randomized trials. *J Natl Cancer Inst,* 103**,** 1799-806.

WAGNER-JOHNSTON, N. D., LINK, B. K., BYRTEK, M., DAWSON, K. L., HAINSWORTH, J., FLOWERS, C. R., FRIEDBERG, J. W. & BARTLETT, N. L. 2015. Outcomes of transformed follicular lymphoma in the modern era: a report from the National LymphoCare Study (NLCS). *Blood,* 126**,** 851-7.

WANG H, AAS E, SMITH A, CROUCH S, ROMAN E & PATMORE R 2016. Forecasting treatment costs of follicular lymphoma: A population-based discrete event simulation. *Value in Health,* 19 (3)**,** A147.

WILD D, WALKER M, PETTENGELL R & LEWIS G 2006. Utility Elicitation in Patients with Follicular Lymphoma. *Value Health,* 9**,** A294.

ZELENETZ, A. D., ABRAMSON, J. S., ADVANI, R. H., ANDREADIS, C. B., BYRD, J. C., CZUCZMAN, M. S., FAYAD, L., FORERO, A., GLENN, M. J., GOCKERMAN, J. P., GORDON, L. I., HARRIS, N. L., HOPPE, R. T., HORWITZ, S. M., KAMINSKI, M. S., KIM, Y. H., LACASCE, A. S., MUGHAL, T. I., NADEMANEE, A., PORCU, P., PRESS, O., PROSNITZ, L., REDDY, N., SMITH, M. R., SOKOL, L., SWINNEN, L., VOSE, J. M., WIERDA, W. G., YAHALOM, J. & YUNUS, F. 2010. NCCN Clinical Practice Guidelines in Oncology: non-Hodgkin's lymphomas. *J Natl Compr Canc Netw,* 8**,** 288-334.

ZHOU, X., WANG, J., ZHANG, J., COPLEY-MERRIMAN, C., TORIGOE, Y., REYES, C., SEYMOUR, J. F., OFFNER, F. C., TRNENY, M. & SALLES, G. A. 2014. Symptoms and toxicity of rituximab maintenance relative to observation following immunochemotherapy in patients with follicular lymphoma. *Hematology,* 20**,** 129-36.

Appendices

The following appendices are provided in a separate file to accompany this submission.

Appendix 1: Draft summary of product characteristics for Gazyvaro

Appendix 2: Draft European Public Assessment Report for Gazyvaro

Appendix 3: Search criteria for clinical SLR

Appendix 4: Studies identified in Clinical Systematic Literature Review

Appendix 5: Search criteria for the systematic literature reviews for the economic model

Appendix 6: Analysis of 9 year follow up data from PRIMA for the economic analysis

Appendix 7: Costs of subsequent treatments based on GALLIUM data

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Single technology appraisal

Obinutuzumab for untreated advanced follicular lymphoma [1020]

Dear Roche,

The Evidence Review Group, Kleijnen Systematic Reviews and the technical team at NICE have now had an opportunity to take a look at the submission received on the 10th May. In general terms they felt that it is well presented and clear. However, the ERG and the NICE technical team would like further clarification relating to the clinical and cost effectiveness data.

Both the ERG and the technical team at NICE will be addressing these issues in their reports.

We request you to provide a written response to this letter to the Institute by **5pm** on **22nd June**. 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. Any supporting documents should be emailed to us separately as attachments or sent on a CD.

If you have any further queries on the technical issues raised in this letter then please contact Anwar Jilani, Technical Lead (Anwar.Jilani@nice.org.uk). Any procedural questions should be addressed to Stephanie Yates, Project Manager (stephanie.yates@nice.org.uk) in the first instance.

Yours sincerely

Nicola Hay
Technical Adviser – Appraisals
Centre for Health Technology Evaluation

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

On behalf of:
Dr Frances Sutcliffe
Associate Director – Appraisals
Centre for Health Technology Evaluation

Section A: Clarification on effectiveness data

Literature searching to inform both clinical and cost effectiveness and utility values

A1. For all searches conducted:

i. Please report the database providers/hosts used to search all databases.

ii. Please provide URLs of conference proceedings, trials registers and organisational websites searches. Please report which search terms and specific years were searched for each conference proceeding.

iii. **Priority question:** For all Medline searches, please check the use of truncation and wildcard/within-word character substitution. The ERG has noted several instances where the truncation or wildcard use has not worked correctly, retrieving incorrect or no results. The search terms are highlighted in yellow.

Please check whether a question mark (?) has been used incorrectly in the place of a truncation character (*). The NLM PubMed database does not support use of a question mark as a wildcard for character substitution. The question mark will be ignored by PubMed and treated as a space or a hyphen, therefore inclusion within a word will not work.

Please examine whether relevant references have been missed as a consequence.

Example 1 (used in Appendix 3, Table 1, line 5, pg 5):{Roche Products Limited, May 2017 [accessed 19.5.17] #35}

Search **naive[TIAB]**	66303
Search **na?ve[TIAB]**	25

Example 2 (used in Appendix 3, Table 1, line 5, pg 5):{Roche Products Limited, May 2017 [accessed 19.5.17] #35}

Search **"newly diagnosed"[Title/Abstract]**	37034
Search **"new*diagnos*"[Title/Abstract]**	0
Search **"newly diagnos*"[Title/Abstract]**	0

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Example 3 (used in Appendix 3, Table 1, line 5, pg 5):{Roche Products Limited, May 2017 [accessed 19.5.17] #35}

Search **((singl*[TIAB] OR doubl*[TIAB] OR treb*[TIAB] OR triple) AND (blind*[TIAB] OR mask*[TIAB]))**	166421
Search **((singl?[TIAB] OR doubl?[TIAB] OR treb?[TIAB] OR triple) AND (blind?[TIAB] OR mask?[TIAB]))**	1132

iv. **Priority question:** Nearly all the searches, with the exception of utility value searches, are restricted to studies that refer to newly diagnosed or untreated patients in the title or abstract. This appears very restrictive as it is possible that a relevant study might not describe line of treatment in the title or abstract. Please clarify why a facet to restrict to line of treatment was included in the searches.

Literature searching – Clinical Effectiveness

A2. Please clarify whether the Embase and Cochrane Library update searches were limited to the publication year range 1998-2016, as reported in Appendix 3, Tables 2 & 3 (pg 7-8).{Roche Products Limited, May 2017 [accessed 19.5.17] #35}

A3. Please confirm where the Cochrane Library search included all databases in the Cochrane Library, or whether the search was restricted to Cochrane Central Register of Controlled Trials (CENTRAL). The numbers reported for this search in the column dated 23.6.15 appear to be the results from the Embase search on the previous page. Please provide numbers for the results from the Cochrane Library or CENTRAL searched on 23.6.15.

i. Following on from the question above, if the Cochrane search was limited to CENTRAL only please explain the rationale for applying a trials study design filter to the search (lines 12-23, page 18-19).{Roche Products Limited, May 2017 [accessed 19.5.17] #35}

ii. If the Cochrane search was not limited to CENTRAL only, please explain the rationale for applying a trials study design filter to the search (lines 12-23, pg 18-19) rather than applying the limit to CENTRAL only.{Roche Products Limited, May 2017 [accessed 19.5.17] #35}

A4. Please check the use of double and single quotation marks used in phrase searching. The incorrect use of quotation marks may have impaired recall within this search.

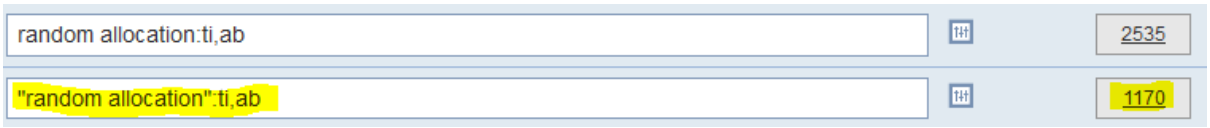
i. Example 1 (used in Appendix 3, Table 3, line 17, pg 7):{Roche Products Limited, May 2017 [accessed 19.5.17] #35}

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

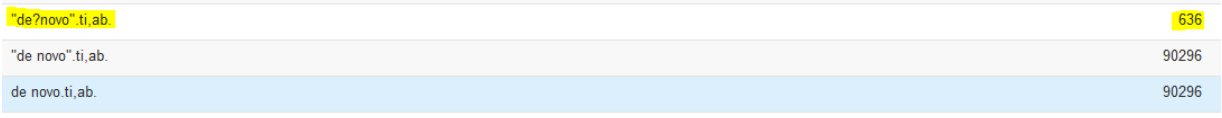
The company's search strategy applied double quotes to this phrase, halving the number of records retrieved. Please examine the implication of this on retrieval of potentially relevant references. Have references been missed as a consequence?



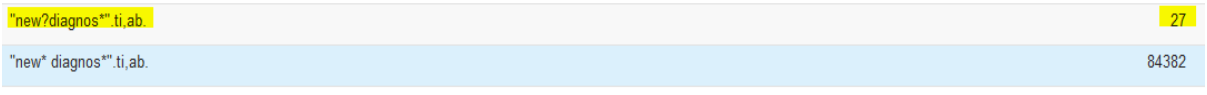
ii. Please clarify why single quotes are used in some lines in the Cochrane strategy (for example lines 2, 4, 16) and double-quotes are used in other lines (lines 5, 15, 17). Do both single and double quotes work in the same way in the database host?

iii. **Priority question:** Please check use of within-word character substitution/wildcard in the Embase and Medline strategies. The ERG has noted several instances where the wildcard use has not worked correctly, retrieving incorrect or no results. The search terms are highlighted in yellow. Please check whether a question mark (?) has been used incorrectly in the place of a truncation character (*). Please examine the implication of this on retrieval of potentially relevant references. Have references been missed as a consequence?

Example 1 (used in Appendix 5, Table 16, line 6, pg 33):{Roche Products Limited, May 2017 [accessed 19.5.17] #35}



Example 2 (used in Appendix 5, Table 16, line 6, pg 33):{Roche Products Limited, May 2017 [accessed 19.5.17] #35}



NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Systematic review - study inclusion

A5. **Priority question:** The cut off for analysis of the GALLIUM trial (including the CSR) was 31 January 2016. Data have been provided from the cut-off on 16 September 2016 'where available'. Please provide a table of the main results including incidence of adverse events from the 16 September 2016 cut-off. In addition, please provide any further data available since the September 2016 cut.

A6. In section 4.1.3 of the company's submission 2 additional exclusion criteria were applied to the systematic review of effectiveness.

A. Please clarify why one of the treatment arms in the included trials had to include rituximab as this would have excluded any trials comparing obinutuzumab and bendamustine regimes.

B. Please clarify whether studies which required patients to have successfully completed induction treatment before entering the maintenance phase were excluded. If these studies were excluded, please explain why.

A7. Please provide a bibliographical list of excluded studies with reasons for exclusion before the additional exclusion criteria were applied, that is the 82 records included in the narrative review in Figure 5 and the 17 records included in the narrative review in Figure 6.

A8. Please explain what is meant by the term 'narrative review' (section 4.1.4) in the context of this submission.

A9. Please provide a bibliographical list of the non-RCTs that were highlighted for the clinical effectiveness review. Was the GAUDI study the only one of relevance to the decision problem of this submission? Did any other studies provide details of adverse events?

A10. For the review of clinical effectiveness, please provide details of the process used for data extraction and assessing methodological quality of the studies (for example whether each of these processes were undertaken by more than one reviewer, did reviewers carry out these tasks independently of each other, whether there was any protocol for identifying and resolving disagreements).

Clinical Effectiveness

A11. **Priority question:** In sections 1.4 and 3.7 of the company's submission it is stated that clinical experts confirmed that the baseline characteristics of the patients with follicular lymphoma in the GALLIUM trial were reflective of the population seen in UK

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

clinical practice. Please provide more details of the clinical experts (full job descriptions and affiliation) and how their opinions were elicited? If surveys were used, please provide the questions and responses. If it was through panel discussions, please provide the transcripts and any notes that were taken during the meeting.

A12. **Priority question:** In the GALLIUM trial 3 types of chemotherapy were combined with obinutuzumab or rituximab. Although the trial was not designed to investigate differences between therapy combinations, differences were noted particularly in adverse event outcomes. The company submission states that there were differing patient characteristics between chemotherapy groups which might explain the results. Please provide the baseline characteristics of participants by type of chemotherapy.

A13. Please confirm the numbers on the flow chart in Figure 9. There appears to be an inconsistency in the numbers who did not start maintenance in the G-chemo arm in GALLIUM.

A14. Please clarify the number of patients in GALLIUM who entered the maintenance phase without successfully completing the induction phase.

A15. How were complete or partial response defined in GALLIUM. Page 60 of the company submission states that a modified version of the Revised Response Criteria was used to ascertain response. How was the Revised Response Criteria modified? Did all patients who started the maintenance phase in GALLIUM have a complete or partial response to therapy?

A16. Section 4.6 of the company's submission assesses the quality of the GALLIUM trial. Although GALLIUM is an open label trial, treatment allocation can still be concealed. Were attempts made to do this? Was the independent review committee (IRC) blind to treatment?

Section B: Clarification on cost-effectiveness data

Literature searching

B1.

i. Please confirm whether the Cochrane Library search included all databases within the Cochrane Library, or whether the search was restricted to NHS Economic Evaluation Database (NHS EED).

ii. Following on from the question above, if the Cochrane search was limited to NHS EED only please explain the rationale for applying an economics filter to the search

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

(lines 12-23, page 18-19).{Roche Products Limited, May 2017 [accessed 19.5.17] #35}

iii. If the Cochrane search was not limited to NHS EED only please explain the rationale for applying an economics filter to the search (lines 12-23, pg 18-19) rather than applying the limit to NHS EED only.{Roche Products Limited, May 2017 [accessed 19.5.17] #35}

B2. Sections on the search strategies for cost-effectiveness are referenced as York, Cochrane or York (adapted). Please provide full references to these sources.

Literature searching - Utility studies search

B3. Sections on the search strategies for utility values are referenced as Sheffield or Cochrane. Please provide full references to these sources.

Literature searching - Resource use

B4. **Priority question:** Please provide the rationale for limiting the Medline and Embase searches to English language publications only. Were any potentially relevant studies excluded on the basis of language?

Progression Free Survival

B5. **Priority question:** The reported hazard ratio for investigator-assessed progression-free survival is 0.66. The hazard ratio for independent review committee progression-free survival (PFS) is 0.71. Data on pages 81 and 82 do not clarify how these analyses were conducted. Therefore, it was not possible to determine why the hazard ratios are different. Please clarify how these hazard ratios were obtained.

B6. **Priority question:** Page 19 of the company submission states that "Investigator-assessed progression-free survival, in line with the primary study endpoint, was extrapolated beyond the observation period in GALLIUM by an exponential distribution, selected by investigating several alternatives modes (i.e., log-normal, log-logistic, Gompertz, generalised gamma or Weibull). This selection was based on the advice of external experts at a UK advisory board on the plausible long-term behaviour, and the observed PFS curves for patients treated with R-chemo+R in the PRIMA study (Salles et al., 2011) and the LymphoCare registry (Nastoupil et al., 2015)". Please provide more details about the UK advisory board (full job descriptions and affiliation of all participants) and how their opinions were elicited? If surveys were used, please provide the questions and responses. If it was through panel discussions, please provide the transcripts and any notes that were taken during the meeting.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

B7. **Priority question:** Please clarify why 9.75 years was assumed as duration of treatment effect on progression-free survival for the base-case. The company submission on page 146 states that "in the PRIMA study there was no indication of a finite duration of treatment effect on PFS in the FL setting, i.e. the proportional hazard assumption for PFS seemed to hold for the entire observation period with longest follow up reaching of up to 9.75 years". However, the PRIMA study did not estimate the relevant treatment effect that is G-chemo+G versus R-chemo+R. Moreover, it states that "clinical advisors suggested that there is no evidence of a finite duration of treatment effect in treatments of FL and that it is plausible that this will be the case for G-chemo+G versus R-chemo+R".

B8. **Priority question:** The proportional hazard assumption does not hold for log-logistic and log-normal models. However, these 2 distributions were considered in sensitivity analyses. Please clarify why these models were considered. Please explain precisely how they were implemented and how the treatment effect was incorporated.

B9. The reason to choose between an exponential or a log-logistic distribution to predict progression free survival is unclear. What was the reason behind the UK advisory board recommending a function representing the mid-range of plausible estimates? Please clarify whether it was based on clinical experience. If so, please provide figures to validate the PFS rates estimated using parametric functions (for example percentage of people surviving progression free at 15 years).

B10. Please indicate why validation against the US LymphoCare registry data was not performed. The ERG acknowledges and understands the limitations of the registry data. However, limitations were also reported for PRIMA and yet it was chosen for validation. Please indicate whether other sources of data for validation are available. If they are available, please provide additional validation exercises as undertaken with the PRIMA data.

B11. Please provide formal statistical tests to further support or reject the choice of proportional hazards.

Transition probability from PFS to death

B12. **Priority question:** Page 147of the company submission states that the "probability of death in PFS was derived from the observed mortality in PFS in the GALLIUM study. Since there were few events, number of deaths and the number of patient-months at risk in PFS were pooled between the arms". This implies that the probability of death in PFS is assumed to be equal for both treatment arms. However, this does not seem to be in line with the figures reported in Table 28 and 29, where the number of deaths observed in the G and R arm are 21 (20.8% of the events) and 14 (9.7% of the events) and 24 (25.8% of the events) and 19 (15.2% of the events),

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

respectively. Therefore, it seems that the number of deaths during PFS is higher in the G arm. Moreover, the number of events reported in Table 66 is 39. It is unclear what the source for this number is, since the number of deaths reported in Table 28 and 29 are 35 and 43, respectively. Furthermore, the number used in the model (sheet 'Death in PFS' cell G10) is not 39 but 38. Please present Table 36 with the correct values and show the number of events, patient-months at risk and monthly rates per treatment arm. Please adjust the model to perform the analysis using different PFS mortality rates for each treatment arm.

Model demographics

B13. Page 77 of the company submission states that the median age in GALLIUM is 59 years. However, page 32 states that the median age of diagnosis in the UK is 65. Please provide a different set of values for use in the model, as shown in Table 60 (that is age, body weight, height, calculated Body Surface Area), where the values shown reflect the characteristics of the advanced FL population in the UK (for example age should be around 65 years).

B14. The proportions of patients in GALLIUM treated with each chemotherapy regimen (CHOP, CVP and bendamustine) are presented in Table 25. In Table 14, these are presented for the general UK population. These are quite different and might indicate that the proportions used in GALLIUM are not reflective of UK clinical practice. Please clarify how the proportions of patients per chemotherapy regimen were used in the model. As an alternative scenario, please present also calculations using the proportions shown in Table 14 instead of those from GALLIUM.

B15. **Priority question**: Please present an additional scenario where the demographic characteristics in the model represent advanced FL population in the UK and concomitant chemotherapy regimens are reflective of UK clinical practice. Please take into account the suggestions made in B13 and B14.

B16. Page 18 of the company submission states that the "study population in GALLIUM is largely reflective of the advanced FL population in the UK. Furthermore, feedback from clinical experts confirms that the baseline characteristics of FL patients enrolled into GALLIUM are reflective of the population seen in UK clinical practice". However, only 21% of the patients in GALLIUM are from the UK. Please provide the arguments used to state that the population in GALLIUM is reflective of the UK population.

Post Progression Survival

B17. Page 149 of the company submission states that the "data was analysed by pooling the treatment arms and stratifying for early and late progression events." Pooling treatment arms can be considered correct if the number of events observed in both

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

arms can be assumed to be the same. These seem not to be reported anywhere. Please provide post-progression survival data per treatment arm and adjust the model to perform the analysis with different progressed disease to death transition probabilities per treatment arm.

B18. Please provide the rationale for the assumption that patients in late progression would require less intensive follow-up when compared to early progressive patients. Please indicate also how “intensive” is defined.

Utility values

B19. **Priority question:** Please provide EQ-5D data (mean, SE and p-values) for both treatment arms in GALLIUM. Please adjust the model to perform the analysis with different utility values per treatment arm.

B20. Please clarify why (not) the utility values should be adjusted for decline in age in the base case.

B21. Throughout the company’s submission it is mentioned that patients in early progressive disease have poorer outcomes than those progressing later. Please clarify whether these “poorer outcomes” refer to mortality only (which was widely discussed) or also refer to health related quality of life. In the latter case, different utility values for early PD and late PD health states should be expected. If applicable, please provide those estimated values.

B22. Only 2 studies were deemed appropriate to source utility values: Wild et al. (conference abstract) and Bec et al. (conference poster). It seems that the main reason for inclusion was that these studies refer to UK data. However, in the base case, GALLIUM data was used, where only 21% of the patients in GALLIUM are from the UK, yet the GALLIUM population was deemed reflective of the UK population. Based on this justification, please indicate whether other (non-UK) studies could be included provided that the population of the study could be considered similar to that in GALLIUM.

Costs and resource use

B23. **Priority question:** Please provide a table presenting costs per cycle (per treatment arm).

B24. **Priority question:** Please provide a full derivation of the administration costs per cycle shown in Table 79.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

B25. Please indicate whether the adverse event rates considered for the cost calculations are also used for the utility values when the disutilities due to adverse events are included in the analysis.

B26. Please clarify whether infusion reactions, premedication, concomitant medication, CT/MRI costs were included in the model. Please indicate whether these were assumed to be the same in both arms, and if so why.

Discontinuation

B27. **Priority question:** Please clarify the differences between the 2 options for treatment discontinuation included in the model.

Cost-effectiveness results

B28. **Priority question:** Please adjust the model to perform the analysis with a longer time horizon (consider a choice where the overall survival [OS] is 0% at the end of the time horizon for all possible extrapolations).

B29. Please provide figures to check the validity of the survival probabilities at the end of the current time horizon (3.8% and 3.3% of the patients are still alive in the treatment and comparator arm, respectively).

B30. Page 188 of the company submission states that “Overall, the predicted OS behaviour seemed plausible and in agreement with observation in GALLIUM. The model seemed to reproduce the observed OS curve in the G-chemo+G arm of GALLIUM but appeared to overestimate (until about 40 months) OS in the R-chemo+R comparator arm”. Please justify this statement by providing the necessary figures. Please explain why the OS behaviour seems plausible and why it appears to overestimate the OS in the comparator arm.

Model implementation

B31. **Priority question:** Please provide plots of PFS Kaplan Meier curves with one parametric distribution at a time to facilitate visual inspection. Please indicate as well the parameterization used in each case; for example, for the exponential distribution this would be $S(t) = \exp(-\lambda t)$, and the source used for the parameterization (for example R, SAS, SPSS, ...).

B32. **Priority question:** Please clarify the differences between the 2 options for “Drug dosing assumption” included in the model and its choice for the base case.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

B33. **Priority question:** Please justify whether vial sharing should be included in the base case. Please indicate the source of the parameter “amount of vial needed to justify its use” and how it is used in the model.

B34. The tornado diagram shown in Figure 37 could not be reproduced. Please confirm or provide the tornado diagram for the base case.

Section C: Textual clarifications and additional points

C1. The method of administration and dosage reported in Table 5 does not completely match with the one presented in Table 2. Please indicate which one is correct.

C2. Page 192 of the company submission states that “Uncertainty was characterised by standard error (if available), covariance matrix or by assuming an error of 20% from the mean if statistical uncertainty was not available”. This is also shown in Table 89. However, in Table 83, 25% is reported. Please indicate which one is correct.

C3. In Table 62 it is mentioned that the probability of remaining in PFS is modelled as a Weibull distribution. However, an Exponential distribution was chosen for the base case. Please indicate which one is correct.

Stephanie Yates
Appraisal Project Manager – Committee C
National Institute for Health and Care Excellence
Level 1A, City Tower
Piccadilly Plaza, Manchester
M1 4BT
By NICE Docs
Manchester

22 June 2017

Re: ID1020 Gazyvaro (obinutuzumab) in combination with chemotherapy for the first-line treatment of patients with advanced follicular lymphoma – Clarification question

Dear Stephanie

Thank you very much for the clarifications questions which we have addressed below. We also have included the revised model addressing the ERG requests and incorporating the updated AE rates.

Regarding the regulatory status, we still anticipate CHMP opinion in July 2017.

We like to point out that the latest anticipated licence wording is:

"Gazyvaro in combination with chemotherapy, followed by Gazyvaro maintenance therapy in patients achieving a response, is indicated for the treatment of patients with previously untreated advanced follicular lymphoma."

The latest version of the draft SmPC is attached in our response as CiC.

Please to not hesitate contacting us for further questions.

Sincerely,

[redacted]

Senior Health Economist

Answers to the clarification questions
Section A: Clarification on effectiveness data

Literature searching to inform both clinical and cost effectiveness and utility values

A1. For all searches conducted:

i. Please report the database providers/hosts used to search all databases.

Please see the table below:

Table 1: Database providers

Database	Database provider
MEDLINE	Pubmed
MEDLINE-IN-PROCESS	Pubmed
EMBASE	Embase.com
Cochrane CENTRAL	Cochrane Library
NHS EED	Cochrane Library

ii. Please provide URLs of conference proceedings, trials registers and organisational websites searches. Please report which search terms and specific years were searched for each conference proceeding.

Please see the enclosed excel file (ID1020 Clarifications Hand searches strategies 2017-06-22 STC noACIC) with the respective tables for information.

iii. **Priority question:** For all Medline searches, please check the use of truncation and wildcard/within-word character substitution. The ERG has noted several instances where the truncation or wildcard use has not worked correctly, retrieving incorrect or no results. The search terms are highlighted in yellow.

Please check whether a question mark (?) has been used incorrectly in the place of a truncation character (*). The NLM PubMed database does not support use of a question mark as a wildcard for character substitution. The question mark will be ignored by PubMed and treated as a space or a hyphen, therefore inclusion within a word will not work.

Please examine whether relevant references have been missed as a consequence.

Example 1 (used in Appendix 3, Table 1, line 5, pg 5):

Search **naive[TIAB]**	66303
Search **na?ve[TIAB]**	25

Example 2 (used in Appendix 3, Table 1, line 5, pg 5):

Search **"newly diagnosed"[Title/Abstract]**	37034
Search **"new*diagnos*"[Title/Abstract]**	0
Search **"newly diagnos*"[Title/Abstract]**	0

Example 3 (used in Appendix 3, Table 1, line 5, pg 5):

Search ((singl*[TIAB] OR doubl*[TIAB] OR treb*[TIAB] OR triple) AND (blind*[TIAB] OR mask*[TIAB]))	166421
Search ((singl?[TIAB] OR doubl?[TIAB] OR treb?[TIAB] OR triple) AND (blind?[TIAB] OR mask?[TIAB]))	1132

We could not fully reproduce the issue as this does only seem to affect searches via PubMed and not via Ovid, for example. However, the following modifications have been made to the search strategy:

- Search terms using * within double quotes have been corrected
- Search terms using ? have been corrected.

The revised search terms are available in a separate Excel file ('ID2020 Clarifications - Revised electronic search strategies 2017 06 22 STC noACIC'). All searches have been re-run, for consistency and new citations have been screened, and full-text reviewed using the same methodology as the one initially used. There was one study (randomised controlled trial) identified from this new search, which was available online on the 24th of March 2017, (1) i.e. after we conducted our search.

As a result, no studies were missed from our searches.

Figure 1: Clinical SLR PRISMA flow chart

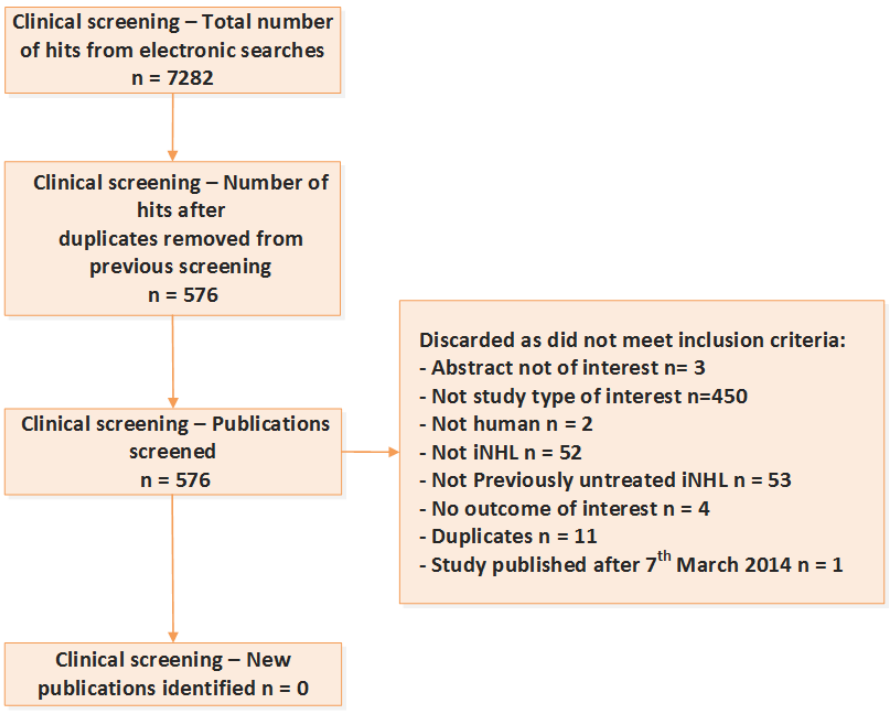


Figure 2: Utility SLR PRISMA flow chart

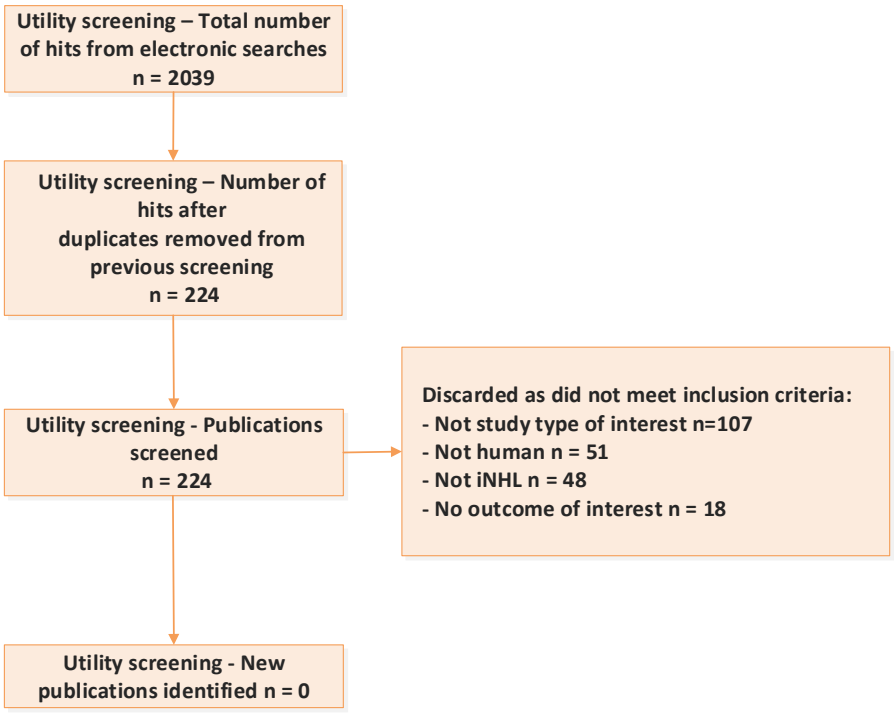
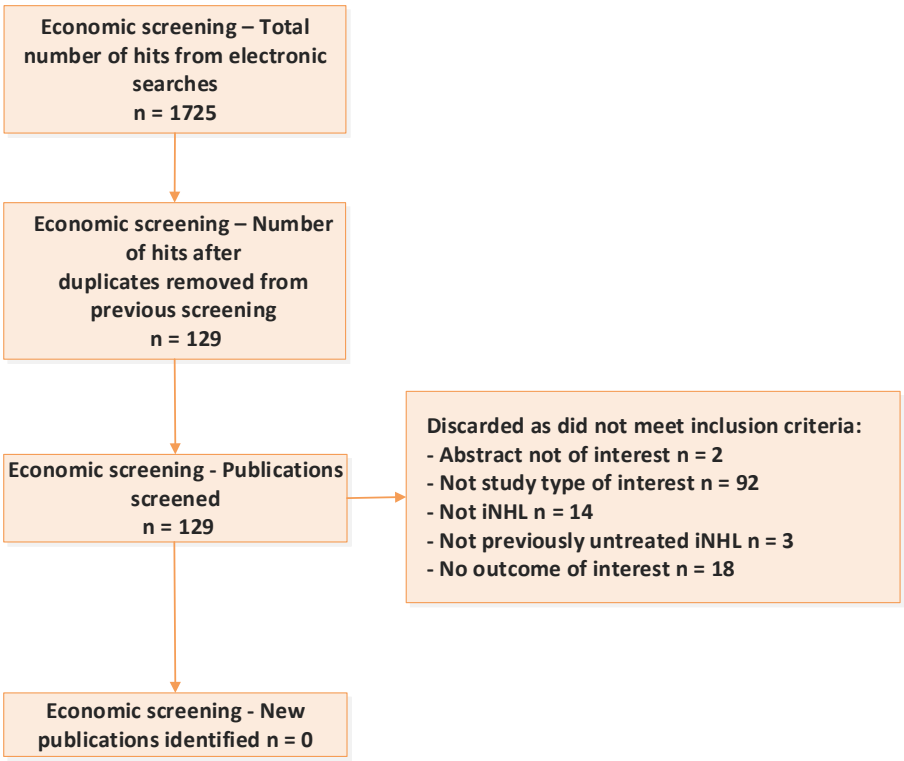


Figure 3: Economic SLR PRISMA flow chart



iv. ***Priority question:*** *Nearly all the searches, with the exception of utility value searches, are restricted to studies that refer to newly diagnosed or untreated patients in the title or abstract. This appears very restrictive as it is possible that a relevant study might not describe line of treatment in the title or abstract. Please clarify why a facet to restrict to line of treatment was included in the searches*.

Line of treatment was restricted at search filter level for the clinical and economic searches in line with the decision problem and the place in therapy for Gazyvaro. In our experience, randomised trials in the follicular lymphoma setting are unlikely not to report on the line of treatment (first line or refractory/relapsed setting) as this is a very important feature of the study design. For resource use, the results of searches performed for this submission showed significant overlap with studies identified in our recent submission for the rituximab-refractory setting in FL (Roche 2016, available at: https://www.nice.org.uk/guidance/gid-ta10020/documents/appraisal-consultation-document-2) indicating that it is very unlikely that relevant studies were missed at filter level.

Literature searching – Clinical Effectiveness

A2. Please clarify whether the Embase and Cochrane Library update searches were limited to the publication year range 1998-2016, as reported in Appendix 3, Tables 2 & 3 (pg 7-8).

All electronic searches have been restricted to 1998 onwards, to match with the launch of rituximab. It is very unlikely that relevant studies were missed prior to 1998. All electronic searches have been updated on the 6th of March 2017. The appendix 3, tables 2&3 do not reflect the latest update of the searches (i.e. this is a typo that do not impact the results).

A3. Please confirm where the Cochrane Library search included all databases in the Cochrane Library, or whether the search was restricted to Cochrane Central Register of Controlled Trials (CENTRAL). The numbers reported for this search in the column dated 23.6.15 appear to be the results from the Embase search on the previous page. Please provide numbers for the results from the Cochrane Library or CENTRAL searched on 23.6.15.

The systematic review of randomised trials was conducted in Cochrane CENTRAL (using the Cochrane Library database provider). The systematic review of economic evaluations was conducted in NHS EED (using the Cochrane Library database provider).

i. Following on from the question above, if the Cochrane search was limited to CENTRAL only please explain the rationale for applying a trials study design filter to the search (lines 12-23, page 18-19).

The search strategy used in the Cochrane Library was incorrect in the submission dossier. No study design filter was used in the electronic search. You will find in the Excel file enclosed the correct search terms that were used. There is no impact on the study selection, since the mistake only appeared in the submission write up.

ii. If the Cochrane search was not limited to CENTRAL only, please explain the rationale for applying a trials study design filter to the search (lines 12-23, pg 18-19) rather than applying the limit to CENTRAL only

See answer to question A3.i

A4. Please check the use of double and single quotation marks used in phrase searching. The incorrect use of quotation marks may have impaired recall within this search.

i. Example 1 (used in Appendix 3, Table 3, line 17, pg 7):

The company's search strategy applied double quotes to this phrase, halving the number of records retrieved. Please examine the implication of this on retrieval of potentially relevant references. Have references been missed as a consequence?

ii. Please clarify why single quotes are used in some lines in the Cochrane strategy (for example lines 2, 4, 16) and double-quotes are used in other lines (lines 5, 15, 17). Do both single and double quotes work in the same way in the database host?

iii. **Priority question:** Please check use of within-word character substitution/wildcard in the Embase and Medline strategies. The ERG has noted several instances where the wildcard use has not worked correctly, retrieving incorrect or no results. The search terms are highlighted in yellow. Please check whether a question mark (?) has been used incorrectly in the place of a truncation character (*). Please examine the implication of this on retrieval of potentially relevant references. Have references been missed as a consequence?

Example 1 (used in Appendix 5, Table 16, line 6, pg 33):

"de?novo".ti,ab.	636
"de novo".ti,ab.	90296
de novo.ti,ab.	90296

Example 2 (used in Appendix 5, Table 16, line 6, pg 33):

"new?diagnos*".ti,ab.	27
"new* diagnos*".ti,ab.	84382

The following modifications have been made to the search strategy:

- Search terms using * within double quotes have been corrected
- Search terms using ? have been corrected.

The revised search terms are available in a separate Excel file. All searches have been re-run, new citations have been screened, and full-text reviewed using the same methodology as the one initially used. There was one study (randomised controlled trial) identified from this new search, which was available online on the 24th of March 2017, (1) i.e. after we conducted our search.

As a result, no studies were missed from our searches.

See question A1 iii for the PRISMA flow charts.

Systematic review - study inclusion

A5. ***Priority question:*** *The cut off for analysis of the GALLIUM trial (including the CSR) was 31 January 2016. Data have been provided from the cut-off on 16 September 2016 'where available'. Please provide a table of the main results including incidence of adverse events from the 16 September 2016 cut-off. In addition, please provide any further data available since the September 2016 cut*.

We would like to point out a textual error in our submission: the correct data for the updated data cut was 10 September 2016 (not 16 September 2016). As highlighted in the submission, an updated CSR and full analysis of this data cut was not available at submission and the detailed results presented in the clinical section were based on the primary analysis with clinical cut-off date of 31 January 2016 and the updated key results from the later data cut (10 September 2017). The analysis of key outcomes also indicated no significant difference to the primary analysis. A comparison of the efficacy data from the primary and updated analyses from GALLIUM is summarised in Table 2 below. A full CSR for the 10 September 2017 clinical cut-off date is now available and enclosed in the reference as CiC.

Table 2: Summary of efficacy data from GALLIUM (primary vs updated analyses – FL population)

	Primary analysis (January 2016 cut-off date)		Updated analysis (September 2016 cut-off date)	
	G-chemo n=601	R-Chemo n=601	G-chemo n=601	R-Chemo n=601
Progression-free survival (INV-assessed, primary endpoint)				
Patients w/ event, n (%)	101 (16.8)	144 (24.0)	120 (20.0)	161 (26.8)
HR (stratified), 95% CI;	0.66 (0.51, 0.85) p=0.0012		0.68 (0.54, 0.87) p=0.0016	
Overall survival				
Patients w/ event, n (%)	35 (5.8%)	46 (7.7%)	43 (7.2%)	52 (8.7%)
HR (stratified), 95% CI	0.75 (0.49, 1.17) p=0.21		0.82 (0.54; 1.22) p=0.32	
Event-Free Survival				
Patients w/ event, n (%)	112 (18.6%)	159 (26.5%)	130 (21.6%)	179 (29.8%)
HR (stratified), 95% CI	0.65 (0.51, 0.83) p=0.0006		0.66 (0.53, 0.83) p=0.0004	
Time to New Anti-Lymphoma Treatment				

Patients w/ event, n (%)	80 (13.3%)	111 (18.5%)	86 (14.3%)	120 (20.0%)
HR (stratified), 95% CI	0.68 (0.51, 0.91) p=0.009		0.68 (0.52, 0.90) p=0.007	
Disease-Free Survival				
Patients included in analysis, n	298	281	307	293
Patients w/ event, n (%)	27 (9.1%)	33 (11.7%)	34 (11.1%)	40 (13.7%)
HR (stratified), 95% CI	0.81 (0.48, 1.35)		0.82 (0.52, 1.31)	
Duration of response				
Patients included in analysis, n	571	567	569	566
Patients w/ event, n (%)	88 (15.4%)	124 (21.9%)	105 (18.5%)	141 (24.9%)
HR (stratified), 95% CI	0.66 (0.50, 0.87)		0.69 (0.53, 0.88)	
Overall response (CR, PR) at end-of-induction				
Without PET n (%)	532 (88.5%)	522 (86.9%)	530 (88.2%)	519 (86.4%)
Δ 95% CI	1.7% (-2.1, 5.5) p=0.33		1.8% (-2.02, 5.68) p=0.30	
With PET n (%)	N=297 255 (85.9%)	N=298 243 (81.5%)	N=297 254 (85.5%)	N=298 242 (81.2%)
Δ 95% CI	4.3% (-1.8,10.4) p=0.19		4.3% (-1.8,10.5) p=0.17	
Complete response at end-of-induction				
Without PET n (%)	117 (19.5%)	143 (23.8%)	112 (18.6%)	145 (24.1%)
Δ 95% CI	-4.3% (-9.1, 0.4) p=0.07		-5.5% (-10.2, -0.78) p=0.02	
With PET n (%)	N=297 185 (62.3%)	N=298 169 (56.7%)	N=297 184 (62.0%)	N=298 169 (56.7%)
Δ 95% CI	5.6% (-2.5, 13.6) p=0.28		5.2% (-2.8, 13.3) p=0.32	

Median follow up primary analysis: 34.5 months; median follow up updated analysis: 41.1 months

Furthermore, all model inputs from GALLIUM in the submission were based on this latest data cut.

The table below summarises the source of AEs reported in the submission document (based on the primary analysis) and the source for AEs in the updated CSR (September 2016 cut-off date). All AEs for the latest data cut (September 2016) have been incorporated in the revised version of the economic model.

Table 3: Source of adverse events reported in submission and CSR summary

Company Submission	Primary CSR	Updated CSR
Table 46	Table 40 (p187) and Table 73 (p266)	Table 35 (p155) and Table 83 (p246)
Table 48	p3124–p3125	p7451–p7452
Table 49	Table 47 (p202)	Table 42 (p171)
Table 50	Table 48 (p203)	Table 43 (p172)
Table 51	Table 51 (p208)	Table 46 (p178)

Please note, after finalisation of the primary and the updated CSRs additional adverse events were identified during source validation; however, these had no impact on the overall adverse event profile of Gazyvaro. Respective reports with the additional AEs for both cut-off dates (January 2016 and September 2016) are included as commercial in confidence references (2, 3).

At this point in time, no data from later data-cuts from the GALLIUM study are available.

[redacted]

A6. *In section 4.1.3 of the company's submission 2 additional exclusion criteria were applied to the systematic review of effectiveness.*

A. Please clarify why one of the treatment arms in the included trials had to include rituximab as this would have excluded any trials comparing obinutuzumab and bendamustine regimes.

One of the treatment arms had to include rituximab as rituximab plus chemotherapy is the standard of care for the first-line treatment of advanced FL. The only study comparing obinutuzumab and bendamustine without rituximab is the GADOLIN study, which is in patients with rituximab-relapsed/refractory FL, therefore not relevant to the indication being appraised.

B. Please clarify whether studies which required patients to have successfully completed induction treatment before entering the maintenance phase were excluded. If these studies were excluded, please explain why.

Studies that required patients to have successfully completed induction treatment before entering the maintenance phase were not excluded.

A7. *Please provide a bibliographical list of excluded studies with reasons for exclusion before the additional exclusion criteria were applied, that is the 82 records included in the narrative review in Figure 5 and the 17 records included in the narrative review in Figure 6.*

Of the 82 records included in the original narrative review, 18 were merged due to multiple publications from the same trial, resulting in 64 studies. In the updated search 11 of the 17 records were merged due to multiple publications, resulting in 6 studies and a total of 70 studies overall. 20 were included in the final narrative review based on the two additional criteria, 50 being excluded. A bibliographical list of these excluded studies can be found below:

Table 4: Citations of excluded studies

Citation	Reason for exclusion
Levy R, J Clin Oncol 2014;32(17):1797-803	Rtx not assessed as treatment
Marschhauser F, J Clin Oncol 2013;31(16):1977-83	Rtx not assessed as treatment
Leblond V, J Clin Oncol 2013;31(3):301-7	Rtx not assessed as treatment
Czuczman MS, Br J Haematol 2012;157(4):438-45	Rtx not assessed as treatment
Lowry L Radiother Oncol 2011;100(1):86-92	Rtx not assessed as treatment
Schuster SJ J Clin Oncol 2011;29(20):2787-94	Rtx not assessed as treatment
Smith SM, Leuk Lymphoma 2009;50(10):1606-17	Rtx not assessed as treatment
Freedman A, J Clin Oncol 2009;;27(18):3036-43	Rtx not assessed as treatment
Nickenig , Cancer 2006;107(5):1014-22	Rtx not assessed as treatment
Sebban C, Blood 2006;108(8):2540-4	Rtx not assessed as treatment
Aviles, A. Med Oncol 2006; 23(2): 295-300	Rtx not assessed as treatment
Hagenbeek A, J Clin Oncol 2006;24(10):1590-6	Rtx not assessed as treatment
Herold M, J Cancer Res Clin Oncol 2006;132(2):105-12	Rtx not assessed as treatment
Aviles, A. Med Oncol 2005; 22(1): 57-62	Rtx not assessed as treatment
Foussard C, Ann Oncol 2005;16(3):466-72	Rtx not assessed as treatment
Lenz, G. Blood 2004; 104(9): 2667-74	Rtx not assessed as treatment
Zinzani P. J Clin Oncol 2004: 22(13): 2654-61	Rtx not assessed as treatment
Ardeshna KM, Lancet 2003;362(9383):516-22	Rtx not assessed as treatment
Haas RL, Ann Hematol 2003;82(7):458-62	Rtx not assessed as treatment
Peterson BA, J Clin Oncol 2003;21(1):5-15	Rtx not assessed as treatment
Aviles A, Eur J Hematol 2002;68(3):144-9	Rtx not assessed as treatment
Rohatiner A, Br J Cancer 2001;85(1):29-35	Rtx not assessed as treatment
Fisher RI, J Clin ONcol 2000;23(33):8447-52	Rtx not assessed as treatment
Kyle RA, Br J Haematol 2000: 108(4): 737-42	Rtx not assessed as treatment
Zinzani P. J Clin Oncol 2000:18(4):773-9	Rtx not assessed as treatment
Coiffier B, Ann Oncol 1999;10(10):1191-7	Rtx not assessed as treatment
Rosenbaum C, ASH 2015 (abstract 2741)	Rtx not assessed as treatment
Saad A. J Clin Oncol 2014: 32(15)	Rtx not assessed as treatment
Gyan E, Blood 2009;113(5):995-1001	Rtx not assessed as treatment
Ha CS, Int J Radiat Oncol Bio Phys 2005;63(1):188-93	Rtx not assessed as treatment
Solal-Celigny P, J Clin Oncol 1998;16(7):2332-8	Rtx not assessed as treatment
Smalley RV, Leukemia 2001;15(7):1118-22	Rtx not assessed as treatment
Hancock, B. Br J Haematol 2009; 144(3): 367-75	Rtx not assessed as treatment
Baldini, L. J Clin Oncol 2003 21(8) 1459-65	Rtx not assessed as treatment
Jones J, ASH 2016, abstract 4388	Rtx not assessed as treatment
Evens A, 2016 ASCO, abstract 7507	Rtx not assessed as treatment; (only in maintenance phase; not reported yet)
Kimby E, Leuk Lymphoma 2015;56(9):2598-607	Conditional on completion of induction
Kahl BS, J Clin Oncol 2014;32(28):3096-102	Conditional on completion of induction
Davies A, Lancet Oncol 2014;15(3):343-52	Conditional on completion of induction
Federico M, J Clin Oncol, 2013;31(12):1506-13	Conditional on completion of induction
Salles G, Lancet 2011;377(9759):42-51	Conditional on completion of induction
Martinelli G, J Clin Oncol 2010;28(29):4480-4	Conditional on completion of induction
Hochster H, J Clin Oncol 2009;27(10):1607-14	Conditional on completion of induction
Buske C, Leukemia 2009;23(1):153-61	Conditional on completion of induction
Ghielmini M , Blood 2004;103(12):4416-23	Conditional on completion of induction
Vitolo U, J Clin Oncol 2013;31(27):3351-9	Conditional on completion of induction
Salar A, J Clin Oncol 2014;32(17):1782-91	Conditional on completion of induction
Lenz, G. J Clin Oncol 2004; 22(24): 4926-33	Conditional on completion of induction
Jurczak, Blood 2016;128:1809	Conditional on completion of induction
Burke JM Blood 2012; 120(21)	Abstract before 2014

A8. *Please explain what is meant by the term 'narrative review' (section 4.1.4) in the context of this submission.*

The systematic literature review was performed to answer specific research questions using a systematic and explicit methodology (i.e. inclusion and exclusion criteria) to identify, select, and critically evaluate results of the studies included in the literature review. These records were included in the narrative review, which comprised a critical review of the findings to determine which studies were relevant to the decision problem.

A9. *Please provide a bibliographical list of the non-RCTs that were highlighted for the clinical effectiveness review. Was the GAUDI study the only one of relevance to the decision problem of this submission? Did any other studies provide details of adverse events?*

Please find the bibliographical list of the non-RCTs in the supporting appendix. The GAUDI study was the only non-RCT identified to be relevant to the decision problem. GAUDI is the only study other than GALLIUM to provide information on adverse events in previously-untreated patients with FL.

A10. *For the review of clinical effectiveness, please provide details of the process used for data extraction and assessing methodological quality of the studies (for example whether each of these processes were undertaken by more than one reviewer, did reviewers carry out these tasks independently of each other, whether there was any protocol for identifying and resolving disagreements).*

As discussed in Section 4.1.3.2 of the Company Submission, all citations identified in the SLR were independently screened by two analysts, with any discrepancies resolved by discussion. A third reviewer was consulted for unresolved disagreements.

Once eligible publications were identified, full papers were obtained and screened again on the basis of the complete manuscript – rather than abstract only – to ensure eligibility. Identical eligibility criteria were used for both steps of the screening processes. As for the first step, two analysts conducted independent reviews of the full publications with a third reviewer consulted for any disagreements.

An independent reviewer undertook the quality check of the data extraction by randomly reviewing 15% of the extracted articles. Any discrepancies were resolved by discussion and a third reviewer was consulted for unresolved disagreements. The 15% QC did not identify any major mistakes, therefore no additional QC was conducted.

Clinical Effectiveness

A11. ***Priority question:*** *In sections 1.4 and 3.7 of the company's submission it is stated that clinical experts confirmed that the baseline characteristics of the patients with follicular lymphoma in the GALLIUM trial were reflective of the population seen in UK clinical practice. Please provide more details of the clinical experts (full job descriptions and affiliation) and how their opinions were elicited? If surveys were used, please provide the questions and responses. If it was through panel discussions, please provide the transcripts and any notes that were taken during the meeting.*

An expert advisory board was consulted at a one-day meeting in April 2017. The panel consisted of the following consultant haematologists specialising in the management of patients with FL, many of whom have experience of obinutuzumab from clinical trials.

Name, professional title	Affiliation
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]

The overall objectives of the meeting were to:

- Understand how previously-untreated, symptomatic patients with advanced FL are currently treated in clinical practice
- Obtain feedback on the clinical efficacy and safety of Gazyvaro in the GALLIUM study and how these data may influence clinical practice
- Gain knowledge on how the GALLIUM data may inform the health economic model for a health technology appraisal.

The supporting appendix provides evidence for the advice obtained from the panel relating to the applicability of the GALLIUM study to UK clinical practice and recommendations for the economic model.

In addition to the advisory board, a draft version of the company submission was sent to [redacted] for her opinion, in which the following response was obtained related to the baseline characteristics of GALLIUM:

"The demographics look standard, other than being younger. The only thing I note is the median time from diagnosis to treatment seemed short (although with a very wide range). If anything, the presumably high number of patients treated very soon after diagnosis suggests selection of a more aggressive cohort."

This feedback was incorporated as part of the Section 4.13 (page 124) of the Company Submission.

A12. *Priority question: In the GALLIUM trial 3 types of chemotherapy were combined with obinutuzumab or rituximab. Although the trial was not designed to investigate differences between therapy combinations, differences were noted particularly in adverse event outcomes. The company submission states that there were differing patient characteristics between chemotherapy groups which might explain the results. Please provide the baseline characteristics of participants by type of chemotherapy.*

The baseline characteristics between chemotherapy subgroups (presented at ICML in June 2017) are summarised below. Overall, high risk patients were more likely to receive CHOP, whereas bendamustine and CVP use was more frequent among older patients and patients with more comorbidity. This reflects the use of chemotherapy regimens in clinical practice.

n (%)	Benda n=686	CHOP n=399	CVP n=117
Median age, years (range)	59 (23–88)	58 (31–85)	59 (32–85)
Age ≥80 years	**23 (3.4)**	3 (0.8)	**4 (3.4)**
Male	332 (48.4)	177 (44.4)	54 (46.2)
Charlson Comorbidity Index score ≥1[†]	**163 (23.8)**	69 (17.3)	22 (18.8)
ECOG PS 2	24 (3.5)	8 (2.0)	6 (5.1)
FLIPI high risk (≥3)	274 (39.9)	**187 (46.9)**	41 (35.0)
Bulky disease (≥7cm)	274 (39.9)	**206 (51.6)**	46 (39.3)

A13. *Please confirm the numbers on the flow chart in Figure 9. There appears to be an inconsistency in the numbers who did not start maintenance in the G-chemo arm in GALLIUM.*

There is a typographical error in Figure 9 of the company submission; of the 557 that completed induction with G-chemo, 18 did not start maintenance (as opposed to the 15 stated in the flow chart). Therefore, 539 patients started maintenance with G.

Furthermore, one additional patient entered the maintenance phase for G without completing the induction phase.

A14. *Please clarify the number of patients in GALLIUM who entered the maintenance phase without successfully completing the induction phase.*

One patient entered the maintenance phase for G without completing the induction phase. All patients who entered the maintenance phase for R had completed the induction phase.

A15. *How were complete or partial response defined in GALLIUM. Page 60 of the company submission states that a modified version of the Revised Response Criteria was used to ascertain response. How was the Revised Response Criteria modified? Did all patients who started the maintenance phase in GALLIUM have a complete or partial response to therapy?*

Complete response or overall response rate (complete or partial response) were defined according to the Revised Response Criteria for Malignant Lymphoma (Cheson BD et al. J Clin Oncol 2007;25(5):579-86). A modified version of the Revised Response Criteria was used to ascertain response; this is summarised in the supporting appendix.

Only patients with complete or partial response at the end-of-induction were to enter the maintenance phase. However, three patients in each arm who had stable disease and one patient in each arm who had progressive disease at the end of induction entered the maintenance phase.

A16. *Section 4.6 of the company's submission assesses the quality of the GALLIUM trial. Although GALLIUM is an open label trial, treatment allocation can still be concealed. Were attempts made to do this? Was the independent review committee (IRC) blind to treatment?*

This was an open-label study; it was not possible to conceal treatment from patients or clinicians due to the differences in dosing schedules, administration rates and premedication between rituximab and obinutuzumab. However, the independent review committee was blinded to treatment.

Section B: Clarification on cost-effectiveness data

Literature searching

B1.

i. Please confirm whether the Cochrane Library search included all databases within the Cochrane Library, or whether the search was restricted to NHS Economic Evaluation Database (NHS EED).

The systematic review of randomised trials was conducted in Cochrane CENTRAL (using the Cochrane Library database provider). The systematic review of economic evaluations was conducted in NHS EED (using the Cochrane Library database provider).

ii. Following on from the question above, if the Cochrane search was limited to NHS EED only please explain the rationale for applying an economics filter to the search (lines 12-23, page 18-19).

The search strategy used in the Cochrane Library was incorrect in the submission dossier. No study design filter was used in the electronic search. You will find in the Excel file enclosed the correct search terms that were used. There is no impact on the study selection, since the mistake only appeared in the submission dossier.

iii. If the Cochrane search was not limited to NHS EED only please explain the rationale for applying an economics filter to the search (lines 12-23, pg 18-19) rather than applying the limit to NHS EED only.

See answer to question B1.ii

B2. Sections on the search strategies for cost-effectiveness are referenced as York, Cochrane or York (adapted). Please provide full references to these sources.

Please find references below:

York adapted:

- http://www.sign.ac.uk/search-filters.html

Cochrane: search terms came from several Cochrane reviews, including the following ones:

- http://www.cochrane.org/CD003805/HAEMATOL_although-the-addition-of-the-anti-cd20-monoclonal-antibody-rituximab-to-chemotherapy-r-chemo-has-been-shown-to-improve-response-rates-and-progression-free-survival-in-patients-with-indolent-or-mantle-cell-lymphoma-the-efficacy-of-r-chemo
- http://www.cochrane.org/CD008909/HAEMATOL_anthracyclines-in-the-treatment-of-follicular-lymphoma-fl-in-adults
- http://www.cochrane.org/CD006552/HAEMATOL_rituximab-as-maintenance-therapy-for-patients-with-follicular-lymphoma
- http://www.cochrane.org/CD004629/HAEMATOL_interferon-alpha-in-the-maintenance-therapy-of-follicular-non-hodgkins-lymphoma

Literature searching - Utility studies search

B3. Sections on the search strategies for utility values are referenced as Sheffield or Cochrane. Please provide full references to these sources.

Sheffield was cited in error, it should be York:

- http://www.indirect-treatment-comparisons.com/wp-content/uploads/2015/06/Poster-374-Sensitivity-Of-A-Search-Filter.pdf

Cochrane: search terms came from several Cochrane reviews, including the following ones:

- http://www.cochrane.org/CD003805/HAEMATOL_although-the-addition-of-the-anti-cd20-monoclonal-antibody-rituximab-to-chemotherapy-r-chemo-has-been-shown-to-improve-response-rates-and-progression-free-survival-in-patients-with-indolent-or-mantle-cell-lymphoma-the-efficacy-of-r-chemo
- http://www.cochrane.org/CD008909/HAEMATOL_anthracyclines-in-the-treatment-of-follicular-lymphoma-fl-in-adults
- http://www.cochrane.org/CD006552/HAEMATOL_rituximab-as-maintenance-therapy-for-patients-with-follicular-lymphoma
- http://www.cochrane.org/CD004629/HAEMATOL_interferon-alpha-in-the-maintenance-therapy-of-follicular-non-hodgkins-lymphoma

Literature searching - Resource use

B4. ***Priority question:*** *Please provide the rationale for limiting the Medline and Embase searches to English language publications only. Were any potentially relevant studies excluded on the basis of language?*

Aim of the resource use literature review was to identify UK studies only as it might be difficult to transfer resource use across countries. Therefore, it was highly unlikely that UK based studies that did not report in English were excluded.

Re-running the original searches showed that 16 non-English references (after deduplication in Medline and Embase) were excluded due to non-English language. On screening of title/abstract none of these citations were found relevant for the review.

In addition we also validated the use of wild card characters (question A1) in the search terms for the resource use SLR. The search was conducted using the Ovid platform and we could not reproduce an issue with the wildcard use. For example) use of ‘na?ve’ (example 1 in A1) always produced more results than ‘naive’. We are therefore confident that the filter for the resource use SLR is appropriate and resulted in identification of relevant UK studies.

Progression Free Survival

B5. ***Priority question:*** *The reported hazard ratio for investigator-assessed progression-free survival is 0.66. The hazard ratio for independent review committee progression-free survival (PFS) is 0.71. Data on pages 81 and 82 do not clarify how these analyses were conducted. Therefore, it was not possible to determine why the hazard ratios are different. Please clarify how these hazard ratios were obtained.*

The estimates for Hazard ratios for investigator assessed PFS as well as independently assessed PFS (PFS-IRC) were derived using a stratified Cox proportional hazards analysis method. Ties in the failure times were handled with approximated likelihood from Efron (4). Analyses were performed using SAS PHREG

procedure. The same analyses were used for the primary analysis (31 January 2016 cut –off date) and the updated analysis (10 September 2016 cut-off date).

The point estimates between the HR therefore differed due to the difference in the underlying PFS events and differences in assessment of progression by investigators or the IRC. However, the differences in hazard ratios are within the statistical uncertainty.

B6. ***Priority question:*** *Page 19 of the company submission states that "Investigator-assessed progression-free survival, in line with the primary study endpoint, was extrapolated beyond the observation period in GALLIUM by an exponential distribution, selected by investigating several alternatives modes (i.e., log-normal, log-logistic, Gompertz, generalised gamma or Weibull). This selection was based on the advice of external experts at a UK advisory board on the plausible long-term behaviour, and the observed PFS curves for patients treated with R-chemo+R in the PRIMA study (Salles et al., 2011) and the LymphoCare registry (Nastoupil et al., 2015)". Please provide more details about the UK advisory board (full job descriptions and affiliation of all participants) and how their opinions were elicited? If surveys were used, please provide the questions and responses. If it was through panel discussions, please provide the transcripts and any notes that were taken during the meeting.*

Questions relating to current clinical practice and the model assumptions were discussed in an advisory board (panel discussion) (see A11 & Appendix B). Regarding the PFS extrapolation, advisors were presented with a graph (according to Figure 25 in the submission) showing the PFS extrapolation the R-chemo+R arm in GALLIUM, representing the current standard of care. Clinical experience seemed to suggest that approximately 60-70% of patients would relapse within 10 years and that therefore an exponential or a log-logistic distribution may be the appropriate PFS distribution choice.

B7. ***Priority question:*** *Please clarify why 9.75 years was assumed as duration of treatment effect on progression-free survival for the base-case. The company submission on page 146 states that "in the PRIMA study there was no indication of a finite duration of treatment effect on PFS in the FL setting, i.e. the proportional hazard assumption for PFS seemed to hold for the entire observation period with longest follow up reaching of up to 9.75 years". However, the PRIMA study did not estimate the relevant treatment effect that is G-chemo+G versus R-chemo+R. Moreover, it states that "clinical advisors suggested that there is no evidence of a finite duration of treatment effect in treatments of FL and that it is plausible that this will be the case for G-chemo+G versus R-chemo+R*".

To our knowledge, there is no indication in the literature of a finite treatment effect of interventions in first line follicular lymphoma. However, this experience is based on rituximab based treatments – either in induction or as maintenance, i.e. comparing R-chemo versus chemo or R maintenance versus observation after induction. Due to the indolent nature of the disease, long-term follow up data is limited in the first line FL setting. The PRIMA study presents a data source with now significantly longer follow up than the GALLIUM study. As the mechanism of action of Gazyvaro as

antiCD20 antibody is similar to that of rituximab, it is expected that the long term effects of treatment observed with rituximab apply to obinutuzumab as well. Gazyvaro has also demonstrated longer term treatment effect versus rituximab in the treatment of chronic lymphocytic leukaemia (CLL). There appears to be no evidence of a finite duration of treatment effect in the CLL11 study that compared G-chlorambucil versus R-chlorambucil with follow up significantly beyond the initial induction treatment phase and median PFS (5).

As we mentioned in the submission, previous economic analyses of rituximab have not assumed an explicit duration of treatment effect in the base case (6). Therefore, the assumption of a finite effect on PFS in our submission is conservative. The treatment effect assumed in the model base case is 9 years, based on the longest observation time of 9.75 years from start of induction in PRIMA and approximately 9.25 years from randomisation to maintenance or observation.

B8. ***Priority question:*** *The proportional hazard assumption does not hold for log-logistic and log-normal models. However, these 2 distributions were considered in sensitivity analyses. Please clarify why these models were considered. Please explain precisely how they were implemented and how the treatment effect was incorporated*.

Log-logistic and Log-normal distributions were investigates as standard as recommended in the NICE DSU methods (7). Please refer to answers B31 below on details of the implementation of the parametric functions in Table 13. The treatment effect for these models was implemented as per the formulas in Table 5 below.

Table 5: Parameter implementation for Log-Logistic and Log-Normal models

Parameters	G-chemo+G	R-chemo+R	Intercept (I)	Treatment (T)	Scale (S)
Log-Normal	μ =I +T; σ=S	μ =I; σ=S	4.948	-0.393	1.618
Log-Logistic	λ=EXP(-(I+T)/S); γ=1/S	λ =EXP(-I/S); γ=1/S	4.758	-0.345	0.752

B9. *The reason to choose between an exponential or a log-logistic distribution to predict progression free survival is unclear. What was the reason behind the UK advisory board recommending a function representing the mid-range of plausible estimates? Please clarify whether it was based on clinical experience. If so, please provide figures to validate the PFS rates estimated using parametric functions (for example percentage of people surviving progression free at 15 years).*

Clinical experts suggested that approximately 60-70% of patients may relapse within 10 years and that therefore, on inspection of the potential PFS extrapolation curves, an Exponential or a Log-logistic function may be the appropriate PFS extrapolation choice (see B6). However, also with the available external PRIMA data, it was not possible to choose between Exponential versus the Log-logistic function. In this situation, we selected the Exponential function as a reasonable choice for the base-case due to the following reasons:

1. In the long term, the rate of progression or death predicted by the Log-logistic function alone would be lower than the general population background mortality

from approximately 20 years onwards. To avoid this inconsistency, PFS is adjusted for background mortality in the model as described in the submission. The Exponential function avoids this problem for a longer extrapolation period, with the rate of progression or death predicted by the Exponential function alone exceeding general background mortality for up to 28 years of extrapolation.

2. The exponential function resulted in more conservative estimates for the PFS benefit and QALYs gained compared to the Log-logistic function as shown in the sensitivity analysis (Table 90 in the submission).

B10. *Please indicate why validation against the US LymphoCare registry data was not performed. The ERG acknowledges and understands the limitations of the registry data. However, limitations were also reported for PRIMA and yet it was chosen for validation. Please indicate whether other sources of data for validation are available. If they are available, please provide additional validation exercises as undertaken with the PRIMA data*.

The main limitation of the LymphoCare R-chemo cohort in Nastoupil et al. for external validation of PFS extrapolation of the R-chemo+R arm, as mentioned in the submission, relates to the use of maintenance in this cohort.

Nastoupil et al. (8) reported 45% and 61% of all patients starting induction with R-CHOP or R-CVP receiving maintenance, respectively. This is considerably less than the 85% observed in PRIMA or approximately 90% in UK clinical practice. This issue may relate to the US maintenance label (see below) and the general problem that any registry reporting long term follow up data will lag behind the current standard of care – i.e. the long-term outcomes are those of a cohort enrolled potentially a decade ago when maintenance was less commonly used. An alternative source may therefore be the publication by Nastoupil et al. (9) that looked specifically at the cohort receiving maintenance. This study seems to indicate higher PFS rates for patients receiving maintenance versus thoe who did not and indicated approximately 60% of patients in PFS after 7 years of follow up on maintenance.

However, an additional limitation is that the US label for rituximab maintenance differs from the EU label as it allows for a different maintenance schedule that is not in agreement with GALLIUM or the EU label (i.e. administration once every two months for up to two years or progression): after CVP induction, in responding patients or with stable disease, the US label allows administering rituximab once weekly for 4 doses at 6-month intervals to a maximum of 16 doses (https://www.gene.com/download/pdf/rituxan_prescribing.pdf).

The LymphoCare cohort reported in Nastoupil et al. is therefore less comparable with the R-chemo+R cohort in GALLIUM than the cohort in PRIMA.

In the UK, data from the Haematological Malignancy Research Network (HMRN) (10, 11) may be a potential source for baseline outcomes on R-chemo+R. However, we are not aware of a publication reporting outcomes for a cohort treated with R-chemo

(followed by R maintenance) and it is likely that long-term outcomes in this registry will lag behind the current standard of care in a similar way as LymphoCare.

B11. *Please provide formal statistical tests to further support or reject the choice of proportional hazards*.

A time-dependent covariate methodology was used to formally test the proportional hazards assumption as recommended by Klein, John P., and Melvin L. Moeschberger. *Survival analysis: techniques for censored and truncated data*. Springer Science & Business Media, 2005. To test the proportional hazards assumption of the treatment effect we artificially created a time dependent covariate, $Z_2(t)$, defined as $Z_2(t)= Z_1\ ln(t\ )$, where Z_1 is an indicator variable for randomized treatment category(0=R-Chemo,1=G-Chemo) and t is time in months from randomization to progression or censoring. A proportional hazards model was fitted to Z_1 and $Z_2(t)$ and the estimates of β_1 and β_2 along with the local test of the null hypothesis that $\beta_2=0$ were obtained. Under this proportional hazards model, the hazard rate at time t is $h(t|Z_1)=ho(t)\exp[\beta1Z_1+\beta2(Z_1\ln(t))]$ so when we compare two individuals, one from G-chemo group and one from R-chemo group the ratio of their hazard rates would equal with
$h[t \mid Z_1 =1]/h[t \mid Z_1 =0]= \exp\{ \beta1+ \beta2*\ln(t) \}$, which depends on t if β_2 is not equal to zero. Thus, a test of $Ho : \beta_2 =0$ is a test for the proportional hazards assumption. The obtained parameter estimate for β_2 was 0.19869 with SE of 0.15055. Wald chi-squared statistics for testing the local hypothesis $\beta_2 =0$ gives a p-value of 0.1869 which support the choice of proportional hazards model.

Transition probability from PFS to death

B12. ***Priority question:*** *Page 147of the company submission states that the "probability of death in PFS was derived from the observed mortality in PFS in the GALLIUM study. Since there were few events, number of deaths and the number of patient-months at risk in PFS were pooled between the arms". This implies that the probability of death in PFS is assumed to be equal for both treatment arms. However, this does not seem to be in line with the figures reported in Table 28 and 29, where the number of deaths observed in the G and R arm are 21 (20.8% of the events) and 14 (9.7% of the events) and 24 (25.8% of the events) and 19 (15.2% of the events), respectively. Therefore, it seems that the number of deaths during PFS is higher in the G arm. Moreover, the number of events reported in Table 66 is 39. It is unclear what the source for this number is, since the number of deaths reported in Table 28 and 29 are 35 and 43, respectively. Furthermore, the number used in the model (sheet 'Death in PFS' cell G10) is not 39 but 38. Please present Table 36 with the correct values and show the number of events, patient-months at risk and monthly rates per treatment arm. Please adjust the model to perform the analysis using different PFS mortality rates for each treatment arm.*

The base case assumed equal probability of death in PFS and was derived by pooling the deaths in PFS in both arms due to the small number of events and the difference not being statistically significant. The deaths contributing to PFS events

(e.g. as reported in the primary CSR in Table 28 and 29 for the January 2016 data cut) were inspected for the reported cause of death and one death with a reason of 'progressive disease' was accounted for in the post progression mortality instead. The PFS death events, deaths not due progressive disease in PFS and patient month at risk for the economic model were based on the September 2016 data cut (12) and are summarised in **Table 6** below.

Table 6: PFS death events (GALLIUM, FL ITT, September 2016 cut-of date)

	N	Events	Months at risk	Monthly Rate (95% CI)
Pooled	1202	38	39,519	0.096% (0.070%-0.132%)
G-chemo	601	23	20,389	0.113% (0.075%-0.170%)
R-chemo	601	15	19,130	0.078% (0.047%-0.130%)

In the revised version of the model mortality in PFS and post-progression (see B17) can be treated separately by arm.

Model demographics

B13. Page 77 of the company submission states that the median age in GALLIUM is 59 years. However, page 32 states that the median age of diagnosis in the UK is 65. Please provide a different set of values for use in the model, as shown in Table 60 (that is age, body weight, height, calculated Body Surface Area), where the values shown reflect the characteristics of the advanced FL population in the UK (for example age should be around 65 years).

The reference for the quote on page 32 cites data HMRN and relates to all FL patients at diagnosis, irrespective of treatment or management of patients. This includes therefore patients with less advanced disease that require no active treatment or patients that may only receive palliative care and not R-chemo. The HMRN also reports patient's age and treatment for follicular lymphoma in the years 2004-2012 (10). In this report a median age of patients treated with chemotherapy is reported as 63.7 (range 19.6-98.3). These patients may be more representative for advanced follicular lymphoma. However, the report does not specifically report the age for R-chemo induction. The median age of patients in GALLIUM was 59.0 years (range: 23 to 88 years) treated with R-chemo+R or G-chemo+G (CSR). Therefore, it may be possible that the GALLIUM cohort is slightly younger than the average UK patient treated in 1L FL (see A11). This could be due to reasons discussed in A11 or that older patients were less likely to enrol, e.g. due to additional burden that may be associated with study participation.

We are not aware of literature reporting other demographic variables, e.g. Body Surface Area (BSA), for advanced follicular lymphoma patients treated with R-chemo first line in the UK. However, a recent publication reports BSA for patients treated for a range of cancers (but not haematological) in England as reported in the SACT data base (13). The average for women was 1.74m^2 (95% CI 1.73–1.74) compared 1.95m^2 (95% CI 1.94–1.95) for men. Based on the proportion of 50.6% male patients

in the GALLIUM cohort, the UK average of 1.85m^2 derived from SACT is in close agreement with the 1.86m^2 in the GALLIUM study. It is therefore unlikely that the dosing of rituximab or chemotherapy would be significantly different in clinical practice compared to the GALLIUM trial.

B14. *The proportions of patients in GALLIUM treated with each chemotherapy regimen (CHOP, CVP and bendamustine) are presented in Table 25. In Table 14, these are presented for the general UK population. These are quite different and might indicate that the proportions used in GALLIUM are not reflective of UK clinical practice. Please clarify how the proportions of patients per chemotherapy regimen were used in the model. As an alternative scenario, please present also calculations using the proportions shown in Table 14 instead of those from GALLIUM.*

The proportion of chemotherapy regimens used in the model corresponds to that in the GALLIUM study (Table 25 in the submission). The proportion present in Table 14 is based on a questionnaire based UK sample (Q4 2016 - Q1 2017 Haematology TAMS, Genactis) based on 157 cases reported by 45 clinicians. On the other hand, in the GALLIUM study, 68% of the UK patients in the study where given Benda and 31% CVP, indicating a more preferential use of bendamustine compared to the market research sample. According to discussions in the advisory board, there are local variations in clinical practice with respect to chemotherapy use and therefore, the appropriate representative average use of the three chemotherapy regimens has some uncertainty.

To our knowledge there is no robust method to conduct a scenario analysis with a different proportion of chemotherapy regimens based on the GALLIUM study results. Somehow re-weighting PFS and OS outcomes by chemotherapy would imply that any differences between the outcome in the chemotherapy strata were due to the chemotherapy only and not due to random error or due to differences in patient characteristics. Both assumptions seem not valid as GALLIUM was not powered for individual chemo sub-groups and patients were not randomised to chemotherapies (resulting in potential differences between chemo groups as discussed in A12). The only feasible scenario analysis may therefore be to assume equal clinical outcomes while weighting chemotherapy, administration and AE costs according to an alternative chemotherapy distribution.

B15. ***Priority question**: Please present an additional scenario where the demographic characteristics in the model represent advanced FL population in the UK and concomitant chemotherapy regimens are reflective of UK clinical practice. Please take into account the suggestions made in B13 and B14.*

Please see the appendix with additional scenario results based on the revised model and the points discussed in B13 and B14.

B16. *Page 18 of the company submission states that the "study population in GALLIUM is largely reflective of the advanced FL population in the UK. Furthermore, feedback from clinical experts confirms that the baseline characteristics of FL patients enrolled*

into GALLIUM are reflective of the population seen in UK clinical practice". However, only 21% of the patients in GALLIUM are from the UK. Please provide the arguments used to state that the population in GALLIUM is reflective of the UK population.

Demographic variables from the UK literature were discussed in B13. We are not aware of studies reporting additional baseline characteristics for UK patients receiving 1L treatment with R-chemo. The GALLIUM sample also presented a significant sample of the UK advanced FL population requiring treatment as indicated by the fact of a separate SACT entry for the study (Aggregate Top 10 Regimens by Diagnostic Group, Available at: http://www.chemodataset.nhs.uk/reports/. Accessed May 2017). Furthermore, a 21% proportion in an international study is a significant representation of patients, given the size of the UK population. In addition, we are not aware that clinical practice in terms of requirements for treatment with R-chemo is significantly different between countries as the treatment with R-chemo is established for several years.

Post Progression Survival

B17. *Page 149 of the company submission states that the "data was analysed by pooling the treatment arms and stratifying for early and late progression events." Pooling treatment arms can be considered correct if the number of events observed in both arms can be assumed to be the same. These seem not to be reported anywhere. Please provide post-progression survival data per treatment arm and adjust the model to perform the analysis with different progressed disease to death transition probabilities per treatment arm.*

PPS (including one PFS event identified as death post progression, see B12) was analysed separately by early and late progression. The numbers at risk and events for Early PD are shown in **Table 7** and the PFS KM curved in Figure 4. In late PD, there were [redacted] patients in risk for R-chemo and [redacted] in G-chemo arm, no event was observed in either of the treatment groups. Therefore, treatment arm specific PPS analysis was only performed for early PD.

Figure 4: Early PD PPS KM per arm (FL, ITT)

[redacted]

Transition rates were derived by fitting an exponential model to the PPS curves and are shown in **Table 7** below.

Table 7: PPS – Early PD (GALLIUM, FL ITT, September 2016 cut-of date)

	Number of Patients	Events	Monthly Rate
R-chemo+R	98	39	1.72%
G-chemo+G	57	19.	1.45%
Pooled	155	58	1.61%

Per treatment arm rates were implemented in the model for Early PD only as there no late PD event in the GALLIUM data set. The scenario with per-treatment arm mortality rates can be run by selecting “Per treatment” in F146 in ‘Model Inputs’, please note that this scenario can only be run when GALLIUM as the source for Early PD PPS is selected.

B18. Please provide the rationale for the assumption that patients in late progression would require less intensive follow-up when compared to early progressive patients. Please indicate also how “intensive” is defined.

It was assumed that late progression would require less intensive care as the disease could be assumed to be progressing more slowly and could be re-treated with R-chemo. This was based on clinical advisors who mentioned that they typically see early progressors in the relapsed setting; patients with long remissions do not

require specialist care and are therefore likely to be re-treated with an R-chemo based regimen. The model assumes a monthly cost of supportive care in early PD based on a frequency of follow up visits equal to induction (PFS), whereas the costs in late PD are assumed to be the same as in long term follow up in PFS (Table 81 in the submission). In the absence of detailed data, cost of next anti-lymphoma treatments were assumed to be the same in early and late PD. Sensitivity analyses presented in the submission indicated that ICERs were not very sensitive to the assumptions.

Utility values

B19. ***Priority question:*** *Please provide EQ-5D data (mean, SE and p-values) for both treatment arms in GALLIUM. Please adjust the model to perform the analysis with different utility values per treatment arm.*

EQ-5D data, analysed with a mixed effects model with health states and treatment as categorical effect (as with Table 68 in the submission) and the difference between the two arms is shown in Table 8 below. There was no statistically significant difference between the arms.

Table 8: GALLIUM EQ-5D utility values by state and treatment arm

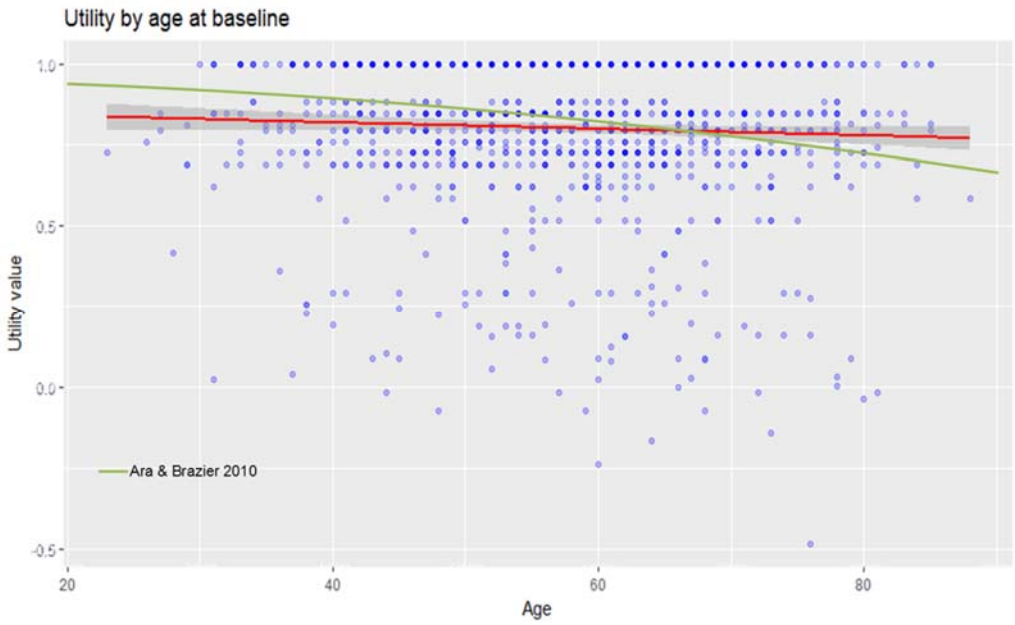
	G-chemo+G		R-chemo+R		Difference	
State	**Estimate**	**Std. Err.**	**Estimate**	**Std. Err.**	**Estimate**	**P-value**
Induction - off tx	0.765	0.032	0.779	0.031	-0.015	0.72
Induction - on tx	0.823	0.015	0.824	0.015	-0.002	0.84
Maintenance & follow-up - off tx	0.826	0.015	0.810	0.015	0.017	0.13
Maintenance & follow-up - on tx	0.834	0.015	0.828	0.014	0.006	0.54
Early progression <= 2yrs	0.767	0.026	0.782	0.022	-0.015	0.62
Late progression > 2yrs	0.820	0.033	0.810	0.030	0.010	0.80

B20. Please clarify why (not) the utility values should be adjusted for decline in age in the base case.

Age effects and the average utility of the general population were reported in Ara and Brazier (14). However, it may only be suitable to use general population values in the absence of disease specific values. In a similar way, it is not obvious that an age

depended decline observed in the general population should translate in the same way to a specific disease. EQ-5D baseline values collected in GALLIUM at baseline (Figure 5, Figure 30 in submission) did not appear to be correlated with age (Pearson correlation: -0.05). Plotting the general population based on Ara and Brazier (14) (with gender proportion from GALLIUM), which appears inconsistent with the observations in GALLIUM in Figure 5. Therefore, baseline utilities in the model were not adjusted by a factor derived from the general population in the base case. Adjustment was performed as a sensitivity analysis.

Figure 5: Baseline utility by age (Figure 30 in submission) versus UK general population



B21. *Throughout the company's submission it is mentioned that patients in early progressive disease have poorer outcomes than those progressing later. Please clarify whether these "poorer outfits comes" refer to mortality only (which was widely discussed) or also refer to health related quality of life. In the latter case, different utility values for early PD and late PD health states should be expected. If applicable, please provide those estimated values.*

This statement refers to the overall survival outcomes. Although it is plausible that early progression is associated with lower utility than late progression, sources of health state utility estimates identified in the SLR have to our knowledge not distinguished whether patients progressed early are late, i.e. between early and late PD.

In the analysis of utility values from the GALLIUM study we were able to distinguish between patients who progressed early compared to those who progressed late. As shown in Table 68 in the submission, average utility values for patients progressing early appear to be lower than those progressing late (and in general higher than figures reported for PD in Wild et al.). However, this may be due to the limited follow up in EQ-5D values beyond the point of progression in GALLIUM leading to more censoring in patients progressing late.

B22. *Only 2 studies were deemed appropriate to source utility values: Wild et al. (conference abstract) and Bec et al. (conference poster). It seems that the main reason for inclusion was that these studies refer to UK data. However, in the base case, GALLIUM data was used, where only 21% of the patients in GALLIUM are from the UK, yet the GALLIUM population was deemed reflective of the UK population. Based on this justification, please indicate whether other (non-UK) studies could be included provided that the population of the study could be considered similar to that in GALLIUM.*

As shown in Table 70 in the submission, studies were deemed less applicable not because of the country setting, but mainly due to other reasons:

- Patients not FL patients or unclear: Levy et al., 2001; Fischbach et al., 2011; Korszun et al., 2014;
- Single centre/small sample size: Friedlich et al., 2006; Andrade-Campos et al., 2014
- Unclear extrapolation from literature: Olin et al., 2010

Wild et al. was UK based but as the additional advantage of a relatively large overall sample in FL patients only.

Costs and resource use

B23. ***Priority question:*** *Please provide a table presenting costs per cycle (per treatment arm).*

The cost per cycle of chemotherapy for administration are in submission Table 79. However, the cycle length for the individual chemotherapies is different. This is accounted for in the model sheet 'dosing calc' where the respective costs are applied to the respective monthly cycle in the model and weighted according the number of patients in each arm and chemotherapy stratum. Furthermore, costs per cycle (and therefore month) differ for the first cycle induction, subsequent induction cycles.

Drug costs were calculated based on the average actual administered dose and the acquisition costs in Table 77 of the submission. These were weighted according the number of patients in each arm and chemotherapy stratum in GALLIUM.

In the maintenance phase for rituximab, weighted costs for MabThera SC or IV were applied to acquisition and administration costs.

The resulting cost schedule is summarised in **Table 9** and **Table 10**.

Table 9: Drug and administration costs per cycle and month - G-chemo+G

				G-benda			G-CHOP			G-CVP			Weighted average		
Days	Cycle - 21 day	Cycle - 28 day	Months	Gazyvaro	Benda	Admin cost	Gazyvaro	CHOP	Admin cost	Gazyvaro	CVP	Admin cost	Gazyvaro	Chemo	Admin
1	1	1	0	[redacted]	93.31	814.54	[redacted]	28.16	430.27	[redacted]	21.84	430.27	[redacted]	64.93	650.86
8	1	1	0	[redacted]	0.00	384.27	[redacted]	0.00	384.27	[redacted]	0.00	384.27	[redacted]	0.00	384.27
15	1	1	0	[redacted]	0.00	384.27	[redacted]	0.00	384.27	[redacted]	0.00	384.27	[redacted]	0.00	384.27
22	2	1	0	[redacted]	0.00	0.00	[redacted]	27.96	430.27	[redacted]	21.66	430.27	[redacted]	11.28	183.28
29	2	2	0	[redacted]	92.61	814.54	[redacted]	0.00	0.00	[redacted]	0.00	0.00	[redacted]	53.16	467.58
43	3	2	1	[redacted]	0.00	0.00	[redacted]	27.68	430.27	[redacted]	21.46	430.27	[redacted]	11.17	183.28
57	3	3	1	[redacted]	92.44	814.54	[redacted]	0.00	0.00	[redacted]	0.00	0.00	[redacted]	53.07	467.58
64	4	3	2	[redacted]	0.00	0.00	[redacted]	27.41	430.27	[redacted]	21.27	430.27	[redacted]	11.06	183.28
85	5	4	2	[redacted]	91.11	814.54	[redacted]	27.08	430.27	[redacted]	21.22	430.27	[redacted]	63.25	650.86
106	6	4	3	[redacted]	0.00	0.00	[redacted]	26.83	430.27	[redacted]	21.23	430.27	[redacted]	10.87	183.28
113	6	5	3	[redacted]	90.54	814.54	[redacted]	0.00	0.00	[redacted]	0.00	0.00	[redacted]	51.97	467.58
127	7	5	4	[redacted]	0.00	0.00	[redacted]	0.00	360.27	[redacted]	21.16	430.27	[redacted]	2.11	160.45
141	7	6	4	[redacted]	89.89	814.54	[redacted]	0.00	0.00	[redacted]	0.00	0.00	[redacted]	51.60	467.58
148	8	6	4	[redacted]	0.00	0.00	[redacted]	0.00	360.27	[redacted]	21.15	430.27	[redacted]	2.11	160.45
Maintenance cycle				[redacted]	0.00	360.27	[redacted]	0.00	360.27	[redacted]	0.00	360.27	[redacted]	0.00	360.27

Table 10: Drug and administration costs per cycle and month - R-chemo+R

				R-benda			R-CHOP			R-CVP			Weighted average		
Days	Cycle - 21 day	Cycle - 28 day	Months	R	Benda	Admin	R	CHOP	Admin	R	CVP	Admin	R	Chemo	Admin
1	1	1	0	[redacted]	92.62	814.54	[redacted]	28.24	430.27	[redacted]	21.58	430.27	[redacted]	64.14	648.30
8	1	1	0	[redacted]	0.00	0.00	[redacted]	0.00	384.27	[redacted]	0.00	0.00	[redacted]	0.00	129.80
15	1	1	0	[redacted]	0.00	0.00	[redacted]	0.00	384.27	[redacted]	0.00	0.00	[redacted]	0.00	129.80
22	2	1	0	[redacted]	0.00	0.00	[redacted]	28.10	430.27	[redacted]	21.38	430.27	[redacted]	11.52	186.14
29	2	2	0	[redacted]	91.88	814.54	[redacted]	0.00	0.00	[redacted]	0.00	0.00	[redacted]	52.13	462.16
43	3	2	1	[redacted]	0.00	0.00	[redacted]	27.83	430.27	[redacted]	21.14	430.27	[redacted]	11.40	186.14
57	3	3	1	[redacted]	91.05	814.54	[redacted]	0.00	0.00	[redacted]	0.00	0.00	[redacted]	51.66	462.16
64	4	3	2	[redacted]	0.00	0.00	[redacted]	27.79	430.27	[redacted]	21.11	430.27	[redacted]	11.39	186.14
85	5	4	2	[redacted]	90.42	814.54	[redacted]	27.48	430.27	[redacted]	21.10	430.27	[redacted]	62.58	648.30
106	6	4	3	[redacted]	0.00	0.00	[redacted]	27.38	430.27	[redacted]	20.94	430.27	[redacted]	11.23	186.14
113	6	5	3	[redacted]	89.11	814.54	[redacted]	0.00	0.00	[redacted]	0.00	0.00	[redacted]	50.56	462.16
127	7	5	4	[redacted]	0.00	0.00	[redacted]	0.00	302.63	[redacted]	21.09	430.27	[redacted]	2.00	143.03
141	7	6	4	[redacted]	88.69	814.54	[redacted]		0.00	[redacted]	0.00	0.00	[redacted]	50.32	462.16
148	8	6	4	[redacted]	0.00	0.00	[redacted]		302.63	[redacted]	21.04	430.27	[redacted]	2.00	143.03
218	11	8	7	[redacted]	0.00	302.63	[redacted]		302.63	[redacted]		302.63	[redacted]	0.00	302.63

B24. ***Priority question:*** *Please provide a full derivation of the administration costs per cycle shown in Table 79*.

Derivation of the administration costs in Table 79 based on the administration schedule and costs in Table 78 is shown in Table 11 below.

Table 11: Administration costs per cycle - derivation of Table 79

Scenario	Tariff	Pharmacy	Transport
1st Cycle G-benda+G:			
Day 1: G + benda	407.00	£11.50	11.77
Day 2: Bedna	361.00	£11.50	11.77
Day 8: G	361.00	£11.50	11.77
Day 15: G	361.00	£11.50	11.77
Cycle Total	1490.00	46.00	47.08
1st Cycle G-CHOP+G, G-CVP+G:			
Day 1: G + CHOP/CVP	407.00	£11.50	11.77
Day 8: G	361.00	£11.50	11.77
Day 15: G	361.00	£11.50	11.77
Cycle Total	£1129.00	£34.50	£35.31
1st and subsequent cycles R-benda+R			
Day 1: R + benda	407.00	£11.50	11.77
Day 2: Bedna	361.00	£11.50	11.77
Cycle Total	£768.00	£23.00	£23.54
1st and subsequent cycles R-CHOP+R, and R-CVP+R	£407.00	£11.50	£11.77
G or R IV maintenance cycle	£337.00	£11.50	£11.77
R SC maintenance cycle	£253.00	-	£11.77

B25. *Please indicate whether the adverse event rates considered for the cost calculations are also used for the utility values when the disutilities due to adverse events are included in the analysis*.

The costs for adverse events (AEs) are always included in the analysis, regardless of whether the disutilities are applied or not as they would need to be accounted for. Disutilities for AEs are only applied in a sensitivity analysis as it is debatable if any influence of the AEs would not have been reflected in the EQ-5D scores collected during treatment already.

B26. Please clarify whether infusion reactions, premedication, concomitant medication, CT/MRI costs were included in the model. Please indicate whether these were assumed to be the same in both arms, and if so why.

Pre-and concomitant medication is assumed to be covered by the respective HRG (DRG) administration costs as per Table 78 in the submission. These medications are not high cost and are expected to be included in the HRG (DRG) costs. Administration costs were assumed to be higher in the G-chemo 1st cycle induction

compared to R-chemo due to the higher number of administration visits (Table 78 and Table 79).

A separate cost to manage infusion reactions (IRR) of grade 3 or 4 was applied under adverse events in the model. The unit costs applied per event were assumed to be £601 (SA31E Malignant Lymphoma, including Hodgkin's and Non-Hodgkin's, with CC Score 2-3). However, this may overestimate the costs as some IRR may be managed during the administration episode and would therefore be included in the administration costs. In the model, the costs are applied on a monthly basis as part of the AE costs for patients on treatment. Monthly costs were calculated based on the number of events per-patient month exposure and were therefore different between the two arms due to the difference in IRR frequency (higher in the G-chemo+G arm compared to R-chemo +R).

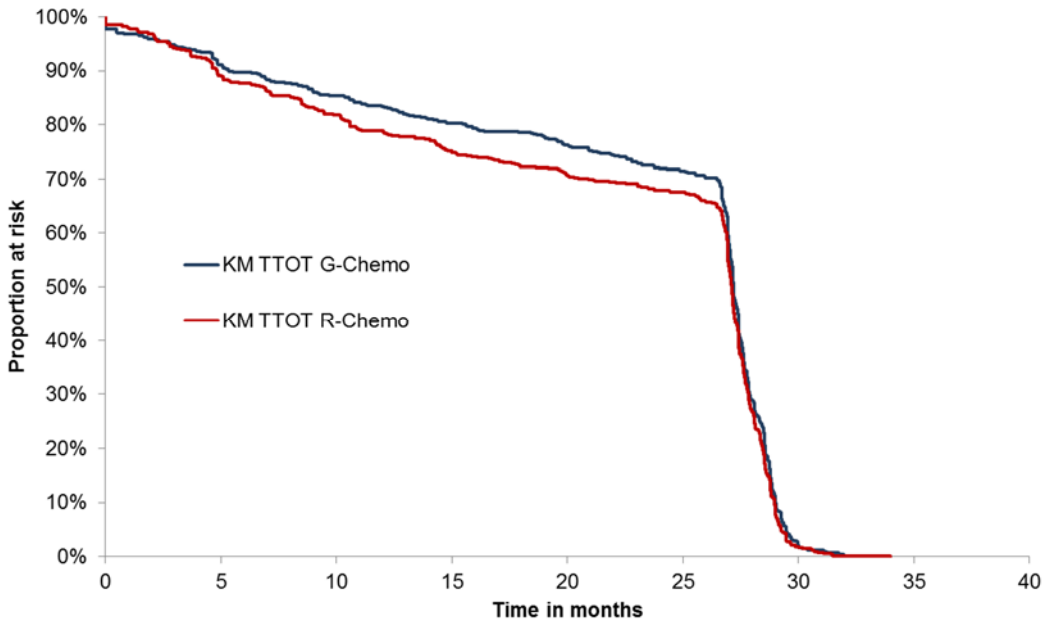
CT/MRI costs were included in the supportive care costs (Table 81 in the submission). With Papaioannou et al. (6), one CT scan in 6 month during induction and one scan in 24 months during maintenance was assumed. There was no reason to assume a difference in supportive care cost per health state between the arms.

Discontinuation

B27. ***Priority question:*** *Please clarify the differences between the 2 options for treatment discontinuation included in the model.*

The base-case option uses the actual observed time on treatment as shown in **Figure 6** below. This includes all discontinuation due to reaching the end of the two year maintenance period, non-response, progression, AEs or other reasons. This corresponds to the actual observed treatment duration associated with the observed efficacy.

Figure 6: GALLIUM time-to-off-treatment KM



The second option assumes treatment until progression, with a maximum of up to two years maintenance, ignoring any discontinuation for other reasons. Efficacy is not adjusted for higher treatment intensity and only drug and administration cost is affected. Although presented in the sensitivity analysis, this scenario is not realistic.

Cost-effectiveness results

B28. ***Priority question:*** *Please adjust the model to perform the analysis with a longer time horizon (consider a choice where the overall survival [OS] is 0% at the end of the time horizon for all possible extrapolations).*

The revised model was adopted to allow a longer time horizon of 50 years. At this stage 0% (0.1%) of patients were expected to be alive in the most optimistic case when assuming a Log-normal PFS function and no finite duration of PFS treatment effect. When updating the model, we also updated the latest UK life tables (ONS 2013-2015 data) available for the general population mortality.

B29. *Please provide figures to check the validity of the survival probabilities at the end of the current time horizon (3.8% and 3.3% of the patients are still alive in the treatment and comparator arm, respectively).*

The data can only be validated against general population life tables: after 40 years a general UK population cohort of matched age is expected to have 5.0.% of survivors (sheet 'Life tables' in the model). The model mortality at this stage is the same in both arms and equal to the general population mortality (2.6% monthly mortality).

B30. *Page 188 of the company submission states that "Overall, the predicted OS behaviour seemed plausible and in agreement with observation in GALLIUM. The model seemed to reproduce the observed OS curve in the G-chemo+G arm of GALLIUM but appeared to overestimate (until about 40 months) OS in the R-chemo+R comparator arm". Please justify this statement by providing the necessary figures. Please explain why the OS behaviour seems plausible and why it appears to overestimate the OS in the comparator arm.*

The model estimates for 12, 24, 36 and 48 months are shown in the table below in comparison the KM estimates. The model appears to overestimate survival in both arms initially, in particular the comparator R-chemo+R arm, and then underestimate survival in both arms. However, the estimates are within the uncertainty of the KM estimates for both arms and therefore model estimates are consistent with observation. Further the model predicts a 20% less OS events at 48 months, i.e. a HR of 0.80 with is consistent with the observed HR of 0.82 (95% CI: 0.54, 1.22) and therefore predicts a plausible difference in OS between the arms.

Table 12: Model OS prediction versus KM estimates

	G-Chemo+G		R-Chemo+R	
Months	Model	KM (95% CI)	Model	KM (95% CI)
12	98.4%	97.8% (96.6%-99.0%)	98.2%	96.4% (94.9%-97.9%)

24	96.0%	95.5% (93.9%-97.2%)	95.1%	93.5% (91.5%-95.5%)
36	93.3%	93.9% (92.0%-95.9%)	91.7%	92.2% (90.0%-94.4%)
48	90.6%	91.5% (88.9%-94.2%)	88.3%	90.6% (88.1%-93.2%)

Model implementation

B31. ***Priority question:*** *Please provide plots of PFS Kaplan Meier curves with one parametric distribution at a time to facilitate visual inspection. Please indicate as well the parameterization used in each case; for example, for the exponential distribution this would be S(t) = exp(-λt), and the source used for the parameterization (for example R, SAS, SPSS, ...).*

Models were fitted to the Kaplan-Meier PFS data using the exponential, Weibull, log-logistic, log-normal, Gompertz and generalized Gamma models presented below. These analyses are specific to the Market access analysis plan and are not part of the Study protocol related Statistical Analysis plan (SAP). The results of such parametric extrapolation were provided as input to the health economics models. To evaluate the goodness of fit of the models, we will use the Akaike Information Criterion (AIC) and diagnostic plots based on transformations of the time scale.

Table 13: Standard parametric models

Model	Survival function	Hazard function
Exponential	$S(t) = \exp(-\lambda t);\ \lambda > 0$	$h(t) = \lambda$
Weibull	$S(t) = \exp(-\lambda t^{\gamma});\ \lambda > 0$ and $\gamma > 0$	$h(t) = \lambda\gamma t^{\gamma-1}$
Log-logistic	$S(t) = \frac{1}{1+\lambda t^{\gamma}};\ \lambda > 0,\ \gamma > 0$	$h(t) = \frac{\lambda\gamma t^{\gamma-1}}{1+\lambda t^{\gamma}}$
Log-normal	$S(t) = 1 - \Phi\left(\frac{\log(t)-\mu}{\sigma}\right);\ \sigma > 0$ where $\Phi(\cdot)$ is standard normal cumulative distribution function	$h(t) = \frac{\frac{1}{\sigma t\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left[\frac{\log(t)-\mu}{\sigma}\right]^2\right)}{1-\Phi\left(\frac{\log(t)-\mu}{\sigma}\right)}$
Gompertz	$S(t) = \exp\left[\frac{\lambda}{\theta}\left(1-e^{\theta t}\right)\right];\ \lambda > 0$	$h(t) = \lambda e^{\theta t}$
Generalized Gamma	$S(t) = \begin{cases} 1 - \Gamma\left[\lambda^{-2}\exp(\lambda[\log(t)-\beta]/\sigma);\ \lambda^{-2}\right] & \text{if } \lambda > 0 \\ \Gamma\left[\lambda^{-2}\exp(\lambda[\log(t)-\beta]/\sigma);\ \lambda^{-2}\right] & \text{if } \lambda < 0 \end{cases}$ where $\Gamma[t;\gamma] = \int_0^t x^{\gamma-1}\exp(-x)dx/\Gamma(\gamma)$	$h(t) = f(t)/S(t)$ where $f(t) = \frac{\lvert\gamma\rvert}{\sigma t\Gamma[\lambda^{-2}]}\left[\lambda^{-2}(e^{-\beta}t)^{\lambda/\sigma}\right]^{\lambda^{-2}}\exp[-\lambda^{-2}(e^{-\beta}t)^{\lambda/\sigma}]$

The models will be fitted using the STEM Macro from the MORSE team (which is based on the SAS procedure LIFEREG for most distributions). Note that the Generalized Gamma model is parameterized differently in the STEM Macro (cf the SAS documentation for PROC LIFEREG for more details).

Plots for the respective fit functions are enclosed as academic in confidence in the file l21223b_PFSINV_FL_plots [AIC].PDF

B32. ***Priority question:*** *Please clarify the differences between the 2 options for "Drug dosing assumption" included in the model and its choice for the base case.*

This option presents a switch between the actual average dose (for each drug and cycle) given to patients in the study versus the planned dose. The base case uses the actual dose rather than the planned does as this might have been altered due to tolerability, for example and corresponds to the actual efficacy observed in the study. As can be seen in the model in the 'dosing calc' sheet, the differences between actual and planned doses are small.

B33. ***Priority question:*** *Please justify whether vial sharing should be included in the base case. Please indicate the source of the parameter "amount of vial needed to justify its use" and how it is used in the model.*

MabThera (rituximab) is the mainstay for 1L treatment of FL as well as other haematological conditions (see SmPC), therefore it can be assumed that most treatment centres have sufficiently high volume of treatment to minimise wastage by vial sharing. In addition, some centres may use dose bands (see for example https://www.england.nhs.uk/commissioning/wp-content/uploads/sites/12/2017/01/national-tables-rituximab-10mgml-v3.pdf) to minimise wastage, by avoiding using a small amount from a new vial, for example. Similar argument may hold for the generic chemotherapy components. For Gazyvaro, vial sharing is not required due to the fixed dosing.

The parameter "amount of vial needed to justify its use" is only available if the option vial sharing = false (no vial sharing) is selected. Selecting 0% results in no use of vial sharing (maximal wastage) and 100% would result in result in 100% use of vial sharing, I,e,, the same scenario as vial sharing.

B34. *The tornado diagram shown in Figure 37 could not be reproduced. Please confirm or provide the tornado diagram for the base case.*

The tornado diagram in the model produces and automated output for continuous variables included in the deterministic sensitivity analysis only. For the tornado diagram in the submission document, additional sensitivity analyses run manually from Table 90 in the submission were included in a manually produced graph (showing values where the difference between upper and lower ICER estimates was >500/QALY).

The source data and graph for the base case are now included in the revised model in the sheet 'Tables Report'. However, these data and graph are not dynamic and will not be updated when a new analysis is run.

Section C: Textual clarifications and additional points

C1. *The method of administration and dosage reported in Table 5 does not completely match with the one presented in Table 2. Please indicate which one is correct*.

Table 2 provides more detailed information on the administration and dosage for Gazyvaro with each chemotherapy regimen, whereas Table 5 is specific for Gazyvaro in combination with bendamustine only. Please find an updated Table 5 below.

Table 14: Standard infusion rate of Gazyvaro in the absence of infusion reactions/hypersensitivity (updated Table 5 from company submission)

Cycle	Day of Treatment	Rate of infusion
Cycle 1	Day 1 (1,000 mg)	Administer at 50 mg/hr. The rate of infusion can be escalated in 50 mg/hr increments every 30 minutes to a maximum of 400 mg/hr.
	Day 8 (1,000 mg)	If no infusion related reaction occurred during the prior infusion when the final infusion rate was 100 mg/hr or faster, infusions can be started at a rate of 100 mg/hr and increased by 100 mg/hr increments every 30 minutes to a maximum of 400 mg/hr.
	Day 15 (1,000 mg)	
Cycles 2–6 (28-day cycle)* or 2–8 (21-day cycle)†	Day 1 (1,000 mg)	
Maintenance	Every two months for two years or until disease progression (whichever occurs first)	

*G-bendamustine
†G-CHOP or G-CVP

C2. *Page 192 of the company submission states that "Uncertainty was characterised by standard error (if available), covariance matrix or by assuming an error of 20% from the mean if statistical uncertainty was not available". This is also shown in Table 89. However, in Table 83, 25% is reported. Please indicate which one is correct.*

The correct value used in the model corresponds to 20%. This affects non-tariff cost values, where variation was not available, such as costs for patient transport, pharmacy or laboratory tests, for example.

C3. *In Table 62 it is mentioned that the probability of remaining in PFS is modelled as a Weibull distribution. However, an Exponential distribution was chosen for the base case. Please indicate which one is correct.*

Exponential is correct as per the base case.

Appendices and enclosed files

Appendix A: Additional scenario analyses and model results

Appendix B: UK advisory board details.

Literature search filter and search strategy information Excel files:

> ‘ID1020 Clarifications Hand searches strategies 2017-06-22 STC noACIC’ & ‘ID2020 ‘Clarifications - Revised electronic search strategies 2017 06 22 STC noACIC’

Draft SmPC: GAZ-EN-VII16-200617-Revision-RtoQ2 [CIC].pdf

Revised model: ID1020 GAZYVARO 1L FL _v1.1 ACIC.xlsb

References

1. Nastoupil LJ, McLaughlin P, Feng L, Neelapu SS, Samaniego F, Hagemeister FB, et al. High ten-year remission rates following rituximab, fludarabine, mitoxantrone and dexamethasone (R-FND) with interferon maintenance in indolent lymphoma: Results of a randomized Study. Br J Haematol. 2017;177(2):263-70.
2. F. Hoffmann-La Roche Ltd. ASSESSMENT OF DELAYED-ENTRY ADVERSE EVENT RECORDS AND UPDATE OF SAFETY RESULTS (CLINICAL CUTOFF: 31 JANUARY 2016). 2017.
3. F. Hoffmann-La Roche Ltd. ASSESSMENT OF DELAYED-ENTRY ADVERSE EVENT RECORDS AND UPDATE OF SAFETY RESULTS (CLINICAL CUTOFF: 10 SEPTEMBER 2016). 2017.
4. Efron B. The Efficiency of Cox's Likelihood Function for Censored Data. Journal of the American Statistical Association. 1977;72(359):557-65.
5. Goede V, Fischer K, Bosch F, Follows G, Frederiksen H, Cuneo A, et al. Updated survival analysis from the CLL11 study: Obinutuzumab versus rituximab in chemoimmunotherapy-treated patients with chronic lymphocytic leukemia. Blood. 2015;126(23):1733.
6. Papaioannou D, Rafia R, Rathbone J, Stevenson M, Buckley Woods H, Stevens J. Rituximab for the first-line treatment of stage III-IV follicular lymphoma (review of Technology Appraisal No. 110): a systematic review and economic evaluation. Health Technol Assess. 2012;16(37):1-253, iii-iv.
7. Latimer NR. NICE DSU technical support document 14: Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. 2013.
8. Nastoupil LJ, Sinha R, Byrtek M, Ziemiecki R, Taylor M, Friedberg JW, et al. Comparison of the effectiveness of frontline chemoimmunotherapy regimens for follicular lymphoma used in the United States. Leuk Lymphoma. 2015;56(5):1295-302.
9. Nastoupil LJ, Sinha R, Byrtek M, Zhou X, Taylor MD, Friedberg JW, et al. The use and effectiveness of rituximab maintenance in patients with follicular lymphoma diagnosed between 2004 and 2007 in the United States. Cancer. 2014;120(12):1830-7.
10. Haematological Malignancy Research Network. Patient's age and treatment for haematological malignancy: a report from the Haematological Malignancy Research Network (HMRN). 2014.
11. Wang H, Aas E, Smith A, Crouch S, Roman E, Patmore R. Forecasting treatment costs of follicular lymphoma: A population-based discrete event simulation. Value in Health. 2016;19 (3):A147.
12. F. Hoffmann-La Roche Ltd. Clinical Study Report, GALLIUM, Report 1075139: updated analysis. 2017.
13. Wallington M. Variations in Body Surface Area of Patients Receiving Chemotherapy Treatment in England. 2013.
14. Ara R, Brazier JE. Using Health State Utility Values from the General Population to Approximate Baselines in Decision Analytic Models when Condition-Specific Data are Not Available. Value in Health. 2011;14(4):539-45.

Appendix A: Additional scenario analyses and model results for clarification answers for Gazyvaro▼ (obinutuzumab) in combination with chemotherapy for the first-line treatment of patients with advanced follicular lymphoma [ID1020]

This appendix presents further scenarios and data in response to the clarification questions. The results of the additional scenarios and base case are based on a revised model version with slightly corrected AE costs including [redacted] relating to the data-cut from [redacted] that formed the basis of the economic analysis for the NICE submission on 10 May 2017. The revised monthly AE costs (submission section 5.5.7) increased from from £53.62 to £56.66 in the G-chemo+G arm and from £45.85 to £48.19 in the R-chemo+R arm, respectively.

Additional changes to the base case was the use of a longer time horizon of 50 years and the use of the latest UK general population life tables (2013-2015 ONS data, sheet 'Life tables') updated when implementing a longer time horizon. All changes were lighlighted in green in the model

The revised base case-results and additional scenarios are pesented below.

Revised Base-case results

The revsed base case results are presented in **Table 1** below and are close to the submission base case.

Table 1: Deterministic base case results (revised Table 85 in submission)

Technologies	Total Costs (£)	Total LYG	Total QALYs	Inc Costs (£)	Inc LYG	Inc QALYs	ICER (£/QALY)
G-chemo+G	[redacted]	13.33	10.01				
R-chemo+R	[redacted]	12.49	9.23	[redacted]	0.84	0.78	[redacted]

Values in the table are discounted and half cycle corrected

Revised disaggregated costs per health state and cost items are summarised in **Table 2** below.

Table 2: Summary of predicted resource use by category of cost (revised Table 88 in submission)

State	Cost (G-chemo)	Cost (R-chemo)	Cost difference	Absolute difference	% of absolute
PFS					
Gazyvaro	[redacted]	0.00	[redacted]	[redacted]	[redacted]
MabThera	0.00	[redacted]	[redacted]	[redacted]	[redacted]
Chemotherapy	370.76	365.43	5.32	5.32	[redacted]

Drug Administration	7,750.79	6,588.61	1,162.18	1,162.18	[redacted]
Adverse Events	1,273.97	1,036.67	237.30	237.30	[redacted]
Supportive Care	7,759.23	6,820.81	938.42	938.42	[redacted]
PFS Total	[redacted]	[redacted]	[redacted]		
Progressive disease					
Supportive care and subsequent treatment costs	10,310.06	11,956.48	-1,646.42	1,646.42	[redacted]
Subsequent treatment costs					
Total PD & PFS	[redacted]	[redacted]	[redacted]	[redacted]	100%

Values in the table are discounted and half cycle corrected

Revised probabilistic sensitivity analysis results

The scatter plot and the corresponding cost-effectiveness acceptability curve for 1,000 simulations are shown in Figure 1 and **Figure 2**, respectively.

Figure 1: Incremental cost and QALY PSA base case results (Figure 35 in submission)

[figure redacted]

Figure 2: Cost-effectiveness acceptability curve (Figure 36 in submission)

[figure redacted]

This analysis indicated that G-chemo+G was more cost-effective than R-chemo+R in [redacted] of simulations at a threshold of £30,000/QALY gained. The probabilistic base-case ICER was £[redacted]QALY.

Revised deterministic sensitivity analysis and scenario analysis results

Revised results of the deterministic sensitivity analysis are shown in Table 3 and the tornado diagram (showing all variables resulting in variation of >£500/QALY in Table 3) in Figure 3. There were no significant differences in the results and conclusion compared to the submission.

Table 3: Deterministic sensitivity analysis for base case (revised Table 90 in submission)

Parameter modified	Base value	High Value*	Low Value*	ICER High	ICER low
Utilities					
Utility in PFS - Induction - On tx	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Utility in PFS - Induction - off tx	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Utility in PFS - Maintenance - off tx	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Utility in PFS - Maintenance - off tx	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Utility in PD - Early progression ≤ 2yrs	0.618	0.693	0.547	[redacted]	[redacted]
Utility in PD - Late progression > 2yrs	0.618	0.693	0.547	[redacted]	[redacted]
Utility source PFS	GALLIUM		Wild		[redacted]

Utility source PD	Wild		GALLIUM		[redacted]
Utility age adjusted	No		Yes		[redacted]
AR Utility included	No		Yes		[redacted]
Costs					
1st administration G-chemo	430	535	347	[redacted]	[redacted]
1st administration R-chemo	430	532	356	[redacted]	[redacted]
Administration G-chemo (subsequent)	384	423	348	[redacted]	[redacted]
Administration R-chemo (subsequent)	384	421	350	[redacted]	[redacted]
Administration maintenece G	360	454	287	[redacted]	[redacted]
Administration maintenece R	303	394	238	[redacted]	[redacted]
Supportive care PFS induction	253	292	223	[redacted]	[redacted]
Supportive care PFS maintenance	83	95	72	[redacted]	[redacted]
Supportive care PFS follow up	58	67	50	[redacted]	[redacted]
AEs - G-chemo+G	54	58	51	[redacted]	[redacted]
AEs - R-chemo+R	46	50	43	[redacted]	[redacted]
Supportive care early PD	231	272	200	[redacted]	[redacted]
Supportive care late PD	58	67	50	[redacted]	[redacted]
Subsequent treatment early PD	13,427	17,038	10,406	[redacted]	[redacted]
Subsequent treatment late PD	13,427	17,065	10,445	[redacted]	[redacted]
Subsequent treatment early/late PD	13,427		5,437.61		[redacted]
Vial sharing	Yes		No		[redacted]
Time on treatment	Actual treatment duration		According to label		[redacted]
MabThera SC use	[redacted] %	80%	40%	[redacted]	[redacted]
Outcomes					
PFS Parametric distribution function	Exponential		Weibull		[redacted]
PFS Parametric distribution function	Exponential		Log-normal		[redacted]
PFS Parametric distribution function	Exponential		Generalised Gamma		[redacted]
PFS Parametric distribution function	Exponential		Log-logistic		[redacted]

PFS Parametric distribution function	Exponential		Gompertz		[redacted]
PFS data set	Investigator		IRC		[redacted]
PFS treatment effect	9 years	No finite duration	5 years	[redacted]	[redacted]
PPS early PD	GALLIUM	PRIMA			[redacted]
PPS early & late PD pooled	Early/late	PRIMA Pooled	GALLIUM Pooled		[redacted]
Discount rate cost & effect	3.50%		1.5%		[redacted]
Time horizon (years)	50		40		[redacted]

Figure 3: Tornado diagram for base case (revised Figure 37 in submission)

[Figure redacted]

Revised results of the scenario analysis in Table 91 and Table 92 of the submission are shown in Table 4 and Table 5, respectively.

Table 4: Scenario analysis – alternative PFS and PPS assumptions (Table 91)

Technologies	Total Costs (£)	Total LYG	Total QALYs	Inc Costs (£)	Inc LYG	Inc QALYs	ICER (£)
G-chemo+G	[redacted]	13.77	10.07				
R-chemo+R	[redacted]	12.98	9.33	[redacted]	0.79	0.74	[redacted]

Values in the table are discounted and half cycle corrected

Table 5: Scenario analysis – assumption on equal post progression QALY and cost (Table 92)

Technologies	PFS Costs (£)	Total LYG in PFS	Tot QALYs in PFS	Inc Costs (£)	Inc LYG	Inc QALYs	ICER (£)
G-chemo+G	[redacted]	8.78	7.20				
R-chemo+R	[redacted]	7.47	6.13	[redacted]	1.31	1.07	[redacted]

Values in the table are discounted and half cycle corrected

Additional scenario analysis

According to the ERG clarification questions additional scenarios were investigated below.

Older starting age (B13, B15)

The starting age of FL patients receiving chemotherapy in the HMRN database cohort was reported as a median of 63.7 (range 19.6-98.3), 4.7 years older than the median age of patients in GALLIUM of 59.0 years (answer to B13). Although, HMRN only captures a subset of the UK population, influence of age was investigated by increasing the mean age in the model from 57.9 years to 62.6 years (i.e. by the difference in median as the HMRN report did not report mean). The increase in the ICER (**Table 6**) can be attributed to the fact that an older cohort would gain less QALYs due to the reduced life expectancy.

Table 6: Scenario analysis – assumption on older 1L FL treatment starting age

Technologies	Total Costs (£)	Total LYG	Tot QALYs	Inc Costs (£)	Inc LYG	Inc QALYs	ICER (£)
G-chemo+G	[redacted]	12.73	9.63				
R-chemo+R	[redacted]	11.99	8.92	[redacted]	0.74	0.71	[redacted]

Values in the table are discounted and half cycle corrected

Different chemotherapy distribution (B14, B15)

Cost for chemotherapy, administration and adverse events were re-weighted according the distributions in Table 14 in the submission. Based on all patients receiving benda, CHOP or CVP (100%) the weights were: 37.18% for benda, 16.67% CHOP and 46.15% CVP, respectively. This resulted in a smaller cost difference between the R-chemo+R and the G-chemo+G arm in drug acquisition, administration and AE costs, decreasing the ICER (**Table 7**).

Table 7: Scenario analysis – assumption different chemotherapy mix (cost only)

Technologies	Total Costs (£)	Total LYG	Tot QALYs	Inc Costs (£)	Inc LYG	Inc QALYs	ICER (£)
G-chemo+G	[redacted]	13.33	10.01				
R-chemo+R	[redacted]	12.49	9.23	[redacted]	0.84	0.78	[redacted]

Values in the table are discounted and half cycle corrected

PFS mortality and PPS by treatment arm (B12, B17)

As described in response to B12 and B17 in the clarification questions, an option was implemented to model separate mortality in PFS and post-progression between the G-chemo+G and the R-chemo+R arm. However, the event rates were low and there was no statistical significant difference between the arms. Whereas the point estimate for the mortality in PFS in the G-chemo+G arm was higher than in the R-chemo+R arm, this was reversed in post-progression, with a lower mortality estimate in the in the G-chemo+G arm compared to the R-chemo+R arm. This scenario resulted in an ICER (**Table 8**) close to the base-case estimate using pooled mortality rates in PFS and PPS. The overall survival benefit predicted by the model was similar between approaches. As discussed in the response to B30, the overall predicted difference in mortality in the model was consistent with the observed OS HR in GALLIUM. However, due to the indolent nature of the condition the event rate was low and the OS difference was not statistically significant.

It should be noted that PPS by arm was implemented for early progression only as there were no deaths observed in either arm for the late progresses due to censoring. Furthermore, utilities per treatment arm were not implemented due to time constrains and the fact the utility differences in GALLUM (response to B19) were small and not statistically significant.

Table 8: Scenario analysis – PFS and PPS (Early PD) by treatment arm

Technologies	Total Costs (£)	Total LYG	Tot QALYs	Inc Costs (£)	Inc LYG	Inc QALYs	ICER (£)
G-chemo+G	[redacted]	13.38	10.04				
R-chemo+R	[redacted]	12.54	9.26	[redacted]	0.84	0.78	[redacted]

Values in the table are discounted and half cycle corrected

Appendix G - professional organisation submission template

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Obinutuzumab for untreated advanced follicular lymphoma [ID1020]

Thank you for agreeing to make a submission on your organisation's view of the technology and the way it should be used in the NHS.

Healthcare professionals can provide a unique perspective on the technology within the context of current clinical practice which is not typically available from the published literature.

To help you in making your submission, we have provided a template. The questions are there as prompts to guide you. It is not essential that you answer all of them.

Please do not exceed the 8-page limit.

About you

Name of your organisation: The Royal College of Radiologists (RCR)

Your name: [redacted], [redacted]

Links with, or funding from the tobacco industry - please declare any direct or indirect links to, and receipt of funding from the tobacco industry: None

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Obinutuzumab for untreated advanced follicular lymphoma [ID1020]

What is the expected place of the technology in current practice?

How is the condition currently treated in the NHS? Is there significant geographical variation in current practice? Are there differences of opinion between professionals as to what current practice should be? What are the current alternatives (if any) to the technology, and what are their respective advantages and disadvantages?

Are there any subgroups of patients with the condition who have a different prognosis from the typical patient? Are there differences in the capacity of different subgroups to benefit from or to be put at risk by the technology?

In what setting should/could the technology be used – for example, primary or secondary care, specialist clinics? Would there be any requirements for additional professional input (for example, community care, specialist nursing, other healthcare professionals)?

If the technology is already available, is there variation in how it is being used in the NHS? Is it always used within its licensed indications? If not, under what circumstances does this occur?

Please tell us about any relevant **clinical guidelines** and comment on the appropriateness of the methodology used in developing the guideline and the specific evidence that underpinned the various recommendations.

The advantages and disadvantages of the technology

NICE is particularly interested in your views on how the technology, when it becomes available, will compare with current alternatives used in the UK. Will the technology be easier or more difficult to use, and are there any practical implications (for example, concomitant treatments, other additional clinical requirements, patient acceptability/ease of use or the need for additional tests) surrounding its future use?

If appropriate, please give your view on the nature of any rules, informal or formal, for starting and stopping the use of the technology; this might include any requirements for additional testing to identify appropriate subgroups for treatment or to assess response and the potential for discontinuation.

If you are familiar with the evidence base for the technology, please comment on whether the use of the technology under clinical trial conditions reflects that observed in clinical practice. Do the circumstances in which the trials were conducted reflect

current UK practice, and if not, how could the results be extrapolated to a UK setting? What, in your view, are the most important outcomes, and were they measured in the trials? If surrogate measures of outcome were used, do they adequately predict long-term outcomes?

What is the relative significance of any side effects or adverse reactions? In what ways do these affect the management of the condition and the patient's quality of life? Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently during routine clinical practice?

Apart from patients with stage I follicular lymphoma for which radical radiotherapy is a curative option, most have advanced stage disease at presentation and are incurable. The median progression-free and overall survival for these patients is 6 to 8 years and 12 to 15 years respectively. Rituximab-chemotherapy induction followed by 2 yrs of maintenance rituximab is the standard first line treatment strategy, and quality of life and time to next treatment are important considerations for patients and clinicians. A higher risk group of patients with FL (approx 30%) relapse within 3 years of frontline treatment but clinical prognostic factors cannot easiy identify these patients who need a more effective treatment strategy. NICE guidelines recommend high dose chemotherapy and an autologous stem cell transplant in second remission, but improvements are needed in frontline therapy.

The Gallium study tested the use of obinutuzumab against rituximab in combination with chemotherapy upfront and as maintenance. 3-year PFS rates were 80% for obinutuzumab and 73.3% for rituximab. There was no difference in 3-year overall survival for the two groups, and response rates were also similar. There were slightly more grade 3+ adverse effects in the obinutuzumab group, and rather surprisingly and unexpectedly, more deaths in the bendamustine arm, regardless of the antibody used.

For some patients, the higher risk of infections will be a concern to be set against the small increase in PFS. More data will be required to see the response to 2nd line treatment following O-chemo and O maintenance as well as overall survival data. In the short term, another option for patients is appreciated but is unlikely to significantly change current practice. However, it has been interesting to see results of immunochemotherapy combinations which will inform future discussions with patients. Data for patients with marginal zone lymphoma, although a smaller group, will also be useful.

Any additional sources of evidence

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Obinutuzumab for untreated advanced follicular lymphoma [ID1020]

Can you provide information about any relevant evidence that might not be found by a technology-focused systematic review of the available trial evidence? This could be information on recent and informal unpublished evidence, or information from registries and other nationally coordinated clinical audits. Any such information must include sufficient detail to allow a judgement to be made as to the quality of the evidence and to allow potential sources of bias to be determined.

Implementation issues

The NHS is required by the Department of Health to provide funding and resources for medicines and treatments that have been recommended by NICE technology appraisal guidance. This provision has to be made within 3 months from the date of publication of the guidance.

If the technology is unlikely to be available in sufficient quantity, or the staff and facilities to fulfil the general nature of the guidance cannot be put in place within 3 months, NICE may advise the Department of Health to vary this direction.

Please note that NICE cannot suggest such a variation on the basis of budgetary constraints alone.

How would possible NICE guidance on this technology affect the delivery of care for patients with this condition? Would NHS staff need extra education and training? Would any additional resources be required (for example, facilities or equipment)?

Equality

NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Please let us know if you think that this appraisal:

- could exclude from full consideration any people protected by the equality legislation who fall within the patient population for which [the treatment(s)] is/are/will be licensed;
- could lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population, e.g. by making it more difficult in practice for a specific group to access the technology;
- could lead to recommendations that have any adverse impact on people with a particular disability or disabilities.

Please tell us what evidence should be obtained to enable the Committee to identify and consider such impacts.

Appendix G - professional organisation submission template

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Obinutuzumab for untreated advanced follicular lymphoma [ID1020]

Thank you for agreeing to make a submission on your organisation's view of the technology and the way it should be used in the NHS.

Healthcare professionals can provide a unique perspective on the technology within the context of current clinical practice which is not typically available from the published literature.

To help you in making your submission, we have provided a template. The questions are there as prompts to guide you. It is not essential that you answer all of them.

Please do not exceed the 8-page limit.

About you

Your name: [redacted]

Name of your organisation: NCRI-ACP-RCP

Links with, or funding from the tobacco industry - please declare any direct or indirect links to, and receipt of funding from the tobacco industry: None

Appendix G - professional organisation submission template

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Obinutuzumab for untreated advanced follicular lymphoma [ID1020]

What is the expected place of the technology in current practice?

How is the condition currently treated in the NHS? Is there significant geographical variation in current practice? Are there differences of opinion between professionals as to what current practice should be? What are the current alternatives (if any) to the technology, and what are their respective advantages and disadvantages?

Lymphoma is the fifth most common cancer in the UK, and follicular lymphoma (FL) is one of the most common subtypes with an annual incidence of approximately 3 per 100,000; this equates to about 3000 new cases each year in the UK. The disease is usually widespread at presentation and runs a chronic relapsing course requiring multiple treatment episodes and culminating in therapy resistance and/or large-cell transformation. Initial treatment for advanced-stage FL is usually commenced for symptoms or complications and typically consists of 6-8 cycles of rituximab (R) combined with one of several different chemotherapy regimens. The use of rituximab in this setting is uncontroversial and approved by NICE (TA 243).

Patients who achieve an anatomical complete (CR) or partial (PR) response then have the option of receiving maintenance therapy with R alone with the aim of delaying disease progression (NICE TA226). However, opinion is divided regarding the routine use of rituximab maintenance in this setting for 3 main reasons. First, data from the pivotal PRIMA trial indicates that the benefit of rituximab maintenance (compared to no maintenance) after frontline chemo-immunotherapy occurs during and shortly after the 2-year period of drug administration and consists of a delay in disease progression in only about 1 in 5 patients treated and a delay in the need for further chemotherapy in only about 1 in 10 patients treated. Second, rituximab maintenance in this setting does not prolong survival even with prolonged follow-up. Third, it increases the risk of infection. This was shown in the PRIMA trial, in a large meta-analysis and in a population-based study which also showed an increase in blood transfusion and growth factor usage.

The results of the GALLIUM trial were presented in December 2016. This phase 3 trial compared rituximab + chemotherapy (CHOP, CVP or bendamustine) followed by R maintenance with obinutuzumab + chemotherapy (CHOP, CVP or bendamustine) followed by obinutuzunab maintenance as frontline treatment for advanced-stage follicular lymphoma. Although the study showed a PFS advantage for the obinutuzumab arm, it failed to demonstrate an OS advantage. Furthermore, the PFS curves diverged during the first 12 months and remained parallel thereafter, with an absolute difference in 3-year PFS of only 4% (77.9% vs 81.9%) as assessed by an independent review committee. Compared with rituximab, obinutuzumab was also associated with more grade ≥3 infections (20% vs 15.6% of patients), infusion-related reactions (12.4% vs 6.75 of patients) and second malignancies (4.7% vs 2.7% of patients).

In summary, rituximab + chemotherapy induction is well established as the standard of care for the initial treatment of advanced-stage follicular lymphoma, but rituximab maintenance in this setting is controversial as it

delays early progression in only ~1 in 5 patients treated, does not prolong survival and increases infection. According to the GALLIUM trial, replacing rituximab with obinutuzumab in this setting delays early progression in ~1 in 25 patients treated, does not prolong survival and is associated with more toxicity.

One unexpected observation in the GALLIUM trial (not mentioned in the meeting abstract) was the high death death rate among the 57% of patients who received bendamustine in combination with either rituximab or obinutuzumab. Bendamustine is not approved as frontline treatment for FL but is nevertheless widely used for this indication in combination with rituximab. This observation calls into question the use of bendamustine as a chemotherapy partner for both rituximab and obinutuzumab in this setting.

Are there any subgroups of patients with the condition who have a different prognosis from the typical patient? Are there differences in the capacity of different subgroups to benefit from or to be put at risk by the technology?

FL is a highly variable disease which can be separated into prognostic groups based on clinical features at diagnosis (FLIPI) or prior to initial treatment (FLIPI-2). However, these scoring systems cannot identify which patients are more or less likely to benefit (or come to harm) from specific treatments. Post-induction FDG PET status is a powerful predictor of early progression following chemo-immunotherapy and is being used to stratify patients in the NCRI phase 3 PETReA trial. In this study, patients who achieve a complete metabolic response (CMR) following rituximab-containing chemo-immunotherapy are randomised to rituximab maintenance versus no further treatment as they have a low risk of early progression even without rituximab maintenance. In contrast, patients who fail to achieve a CMR are randomised to rituximab maintenance versus rituximab plus lenalidomide (treatment escalation) as they have a much higher risk of early progression.

In what setting should/could the technology be used – for example, primary or secondary care, specialist clinics? Would there be any requirements for additional professional input (for example, community care, specialist nursing, other healthcare professionals)?

Rituximab-containing chemo-immunotherapy and maintenance therapy for FL are delivered in secondary care (usually in a day-ward) under the supervision of a haemato-oncology or medical oncology team. Replacing IV rituximab with obinutuzumab should not have any significant additional impact on healthcare professional resources. However, it should be noted that many patients nowadays receive the subcutaneous formulation of rituximab as maintenance treatment for FL. SC rituximab has the advantage of being administered over ~5 mins instead of the ~4-6 hours required for IV rituximab and therefore saves a significant amount of day-ward and nursing time, as well as being more convenient for patients. Replacing SC rituximab with obinutuzumab would therefore require significantly more day-ward and nursing time and be less convenient for patients.

If the technology is already available, is there variation in how it is being used in the NHS? Is it always used within its licensed indications? If not, under what circumstances does this occur?

Obinutuzumab is already approved by NICE in combination with chlorambucil as a possible treatment for previously untreated CLL (TA343). It is not routinely used outside of its NICE approved indication.

Please tell us about any relevant **clinical guidelines** and comment on the appropriateness of the methodology used in developing the guideline and the specific evidence that underpinned the various recommendations.

NICE has produced guidance on the diagnosis and management of non-Hodgkin's lymphoma (NG52) which includes a section (1.3) on FL. In the section on advanced-stage symptomatic FL, previous NICE guidance on rituximab as part of initial chemoimmunotherapy (TA243) and post-induction maintenance (TA226) is cited.

The advantages and disadvantages of the technology

NICE is particularly interested in your views on how the technology, when it becomes available, will compare with current alternatives used in the UK. Will the technology be easier or more difficult to use, and are there any practical implications (for example, concomitant treatments, other additional clinical requirements, patient acceptability/ease of use or the need for additional tests) surrounding its future use?

The practicalities of administering obinutuzumab are similar to those of administering IV rituximab, although obinutuzumab is associated with a higher rate of infusion-related reactions.

As previously mentioned, many patients receive the SC formulation of rituximab which is administered over ~5 mins instead of ~4-6 hours. Replacing SC rituximab with obinutuzumab would therefore require significantly more day-ward and nursing time and be less convenient for patients.

Another consideration is the imminent advent of IV rituximab biosimilars which are likely to be priced much lower than the IV Mabthera.

If appropriate, please give your view on the nature of any rules, informal or formal, for starting and stopping the use of the technology; this might include any requirements for additional testing to identify appropriate subgroups for treatment or to assess response and the potential for discontinuation.

As explained earlier, the PETReA trial is testing the hypothesis that patients who achieve a CMR on FDG PET following frontline rituximab-based chemoimmunotherapy do not benefit from rituximab maintenance. The same hypothesis could be extended to obinutuzumab but there is currently insufficient evidence to support the routine use of FDG PET for this purpose.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Obinutuzumab for untreated advanced follicular lymphoma [ID1020]

If you are familiar with the evidence base for the technology, please comment on whether the use of the technology under clinical trial conditions reflects that observed in clinical practice. Do the circumstances in which the trials were conducted reflect current UK practice, and if not, how could the results be extrapolated to a UK setting? What, in your view, are the most important outcomes, and were they measured in the trials? If surrogate measures of outcome were used, do they adequately predict long-term outcomes?

The UK made a major contribution to the GALLIUM trial. There was nothing particularly stringent about the entry criteria, and the profile of patients recruited was as expected. However, as with any trial, it is inevitable that older, frail patients and those with a high burden of co-morbidity were under-represented.

Two aspects of the trial design should be noted. First, the dose of obinutuzumab was significantly higher than that of rituximab. The latter was given at 375 mg/m^2 at each admistration whereas obinutuzumab was given at a flat dose of 1000 mg. Furthermore, obinutuzumab was given on day 1, 8 and 15 of cycle 1 as well as day 1 of each subsequent induction cycle, whereas rituximab was given only once with each induction cycle.

Second, the primary endpoint in GALLIUM was PFS as determined by non-blinded local investigators despite the availability of PFS data as determined by a blinded endpoint review committee. In fact, the difference in PFS at 3 years was 6.7% when determined locally (80% vs 73.3%) but only 4% when determined by the ERC (81.8% vs 77.9%).

What is the relative significance of any side effects or adverse reactions? In what ways do these affect the management of the condition and the patient's quality of life? Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently during routine clinical practice?

The most common grade ≥3 toxicities in the obinutuzumab arm of GALLIUM were neutropenia, febrile neutropenia, thrombocytopenia and infusion-related reactions, all of which occurred more frequently than in the rituximab arm. Grade ≥3 infections and second neoplasms were also more common in the obinutuzumab arm.

Infusion-related reactions can be very distressing for patients and delay drug administrations. Serious infections are not only unpleasant for patients but can also be life-threatening if not promptly treated and usually result in hospitalisation. We have also learned with rituximab that some rare but extremely serious infections (e.g. PML) may occur a long time after drug exposure and may not be captured by the usual SAE reporting system.

Any additional sources of evidence

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Obinutuzumab for untreated advanced follicular lymphoma [ID1020]

Can you provide information about any relevant evidence that might not be found by a technology-focused systematic review of the available trial evidence? This could be information on recent and informal unpublished evidence, or information from registries and other nationally coordinated clinical audits. Any such information must include sufficient detail to allow a judgement to be made as to the quality of the evidence and to allow potential sources of bias to be determined.

Implementation issues

The NHS is required by the Department of Health to provide funding and resources for medicines and treatments that have been recommended by NICE technology appraisal guidance. This provision has to be made within 3 months from the date of publication of the guidance.

If the technology is unlikely to be available in sufficient quantity, or the staff and facilities to fulfil the general nature of the guidance cannot be put in place within 3 months, NICE may advise the Department of Health to vary this direction.

Please note that NICE cannot suggest such a variation on the basis of budgetary constraints alone.

How would possible NICE guidance on this technology affect the delivery of care for patients with this condition? Would NHS staff need extra education and training? Would any additional resources be required (for example, facilities or equipment)?

The availability of ofatumumab for this indication in the NHS would present significant additional resource requirements other than those already mentioned.

Equality
NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Please let us know if you think that this appraisal:

- could exclude from full consideration any people protected by the equality legislation who fall within the patient population for which [the treatment(s)] is/are/will be licensed;
- could lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population, e.g. by making it more difficult in practice for a specific group to access the technology;
- could lead to recommendations that have any adverse impact on people with a particular disability or disabilities.

Please tell us what evidence should be obtained to enable the Committee to identify and consider such impacts.

We do not believe this appraisal raises any issues of equality.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Obinutuzumab for untreated advanced follicular lymphoma [ID1020]

Thank you for agreeing to give us a statement on your view of the technology and the way it should be used in the NHS.

Healthcare professionals can provide a unique perspective on the technology within the context of current clinical practice which is not typically available from the published literature.

To help you in making your statement, we have provided a template. The questions are there as prompts to guide you. It is not essential that you answer all of them.

Please do not exceed the 8-page limit.

About you

Your name: Graham Collins

Name of your organisation: Royal College of Physicians

Are you (tick all that apply):

- a specialist in the treatment of people with the condition for which NICE is considering this technology? - yes
- a specialist in the clinical evidence base that is to support the technology (e.g. involved in clinical trials for the technology)? - yes
- an employee of a healthcare professional organisation that represents clinicians treating the condition for which NICE is considering the technology? If so, what is your position in the organisation where appropriate (e.g. policy officer, trustee, member etc.)? – yes, chair of the British Society of Haeamtology Special Interest Group for Lymphoma
- other? (please specify)

Links with, or funding from the tobacco industry - please declare any direct or indirect links to, and receipt of funding from the tobacco industry: No

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

What is the expected place of the technology in current practice?

How is the condition currently treated in the NHS? Is there significant geographical variation in current practice? Are there differences of opinion between professionals as to what current practice should be? What are the current alternatives (if any) to the technology, and what are their respective advantages and disadvantages?

Front line treatment of advanced symptomatic follicular lymphoma is with rituximab+chemotherapy (bendamustine, CVP or CHOP). Patients then are offered maintenance rituximab two monthly for two years. This is not curative but can result in remission which last on average 6-8 years. Patients relapsing early (within 2 years) have a significantly shortened overall survival compared with those relapsing later. There is variation in the chemotherapy component but not in the antibody component. There is widespread agreement that this is the standard approach. The only alternative is single agent rituximab but this is less effective and the remissions less durable that R+chemo. Currently there are no other anti-CD20 antibodies that are used for this indication.

Are there any subgroups of patients with the condition who have a different prognosis from the typical patient? Are there differences in the capacity of different subgroups to benefit from or to be put at risk by the technology?

The FLIPI-1 and FLIPI-2 scores are clinical risk scores which identify low, medium and high risk patients. The high risk group do have significantly shorter overall survival outcomes than the low and medium. However this does not currently affect the treatment approach. Other risk scores generally apply to first remission duration which is not relevant for the technology under appraisal which is for front line use. There are no biological markers at diagnosis which are used in routine practise to identify risk groups.

In what setting should/could the technology be used – for example, primary or secondary care, specialist clinics? Would there be any requirements for additional professional input (for example, community care, specialist nursing, other healthcare professionals)?

The technology would be used in chemotherapy delivery day treatment units of hospitals. Delivery at home or in other settings in an option but only with good oversight and governance from specialists, and delivery by chemotherapy trained nurses.

If the technology is already available, is there variation in how it is being used in the NHS? Is it always used within its licensed indications? If not, under what circumstances does this occur?

It is currently only available for the treatment of older / more frail patients with CLL in combination with chlorambucil. It is not being used for follicular lymphoma although an appraisal is ongoing evaluating it in combination with bendamustine for rituximab-refractory follicular lymphoma (it has a license for the indication).

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Please tell us about any relevant **clinical guidelines** and comment on the appropriateness of the methodology used in developing the guideline and the specific evidence that underpinned the various recommendations.

There is a NICE guideline on the management of non-Hodgkin lymphoma although the current indication was not covered. ESMO guidelines recommend R+chemo for the treatment of symptomatic advanced stage follicular lymphoma however this was published prior to the GALLIUM study being presented (which is the seminal phase III assessing obinutuzumab for this indication)

The advantages and disadvantages of the technology

NICE is particularly interested in your views on how the technology, when it becomes available, will compare with current alternatives used in the UK. Will the technology be easier or more difficult to use, and are there any practical implications (for example, concomitant treatments, other additional clinical requirements, patient acceptability/ease of use or the need for additional tests) surrounding its future use?

Obinutuzumab does take slightly longer to infuse than rituximab which will have some impact on day unit capacity. There are slightly more infusion reactions (in CLL there are markedly more infusion reactions but this is LESS so for follicular lymphoma as the peripheral white count is less commonly raise and it is a raised white count that predicts for this reaction). Neutropenia and infection rates are slightly higher than for rituximab which will have a modest impact of health care resources.

If appropriate, please give your view on the nature of any rules, informal or formal, for starting and stopping the use of the technology; this might include any requirements for additional testing to identify appropriate subgroups for treatment or to assess response and the potential for discontinuation.

As with all chemotherapy combinations, the treatment should be stopped after 3-4 cycles if there is no response. This is standard practise already. There are no other early stopping rules that would apply. There are no additional tests to identify subgroups.

If you are familiar with the evidence base for the technology, please comment on whether the use of the technology under clinical trial conditions reflects that observed in clinical practice. Do the circumstances in which the trials were conducted reflect current UK practice, and if not, how could the results be extrapolated to a UK setting? What, in your view, are the most important outcomes, and were they measured in the trials? If surrogate measures of outcome were used, do they adequately predict long-term outcomes?

The GALLIUM trial is a straightforward trial. The standard arm very much reflects current UK practice: sites chose between R-benda, R-CVP or R-CHOP. The experimental arm was G-benda, G-CVP or G-CHOP (obinutuzumab has the trade

name Gazyva hence then 'G'). So the trial is applicable to the UK. The most important outcome in a cancer trial is overall survival. However progression free survival is often taken as a surrogate for this. I support this as otherwise the follow up of indolent lymphoma trials would be so long, that no progress in treatment could be made. Also, patient groups frequently emphasise the importance of remission duration (not surprisingly) which is reflected by PFS. Time to next treatment is also an important (although more subjective) outcome measure. This was measured in the trial. The trial showed a statistically significant prolongation of PFS with obinutuzumab which in my view would be clinically relevant for patients.

What is the relative significance of any side effects or adverse reactions? In what ways do these affect the management of the condition and the patient's quality of life? Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently during routine clinical practice?

The trial showed a slight increase in neutropenia and infections. This is important as it could impair the quality of life of patients in remission. However the increase in toxicity was not marked and the adverse events were largely manageable. The benefit: risk would of course need to be discussed with individual patients, but in my view the most important factor for patients is remission duration.

There have been no adverse events that have come to light in the CLL setting, where obinutuzumab is currently being used.

Equality and Diversity

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Please let us know if you think that this appraisal:

- Could exclude from full consideration any people protected by the equality legislation who fall within the patient population for which [the treatment(s)] is/are/will be licensed;
- Could lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population, e.g. by making it more difficult in practice for a specific group to access the technology;
- Could lead to recommendations that have any adverse impact on people with a particular disability or disabilities

Please tell us what evidence should be obtained to enable the Committee to identify and consider such impacts

Any additional sources of evidence

Can you provide information about any relevant evidence that might not be found by a technology-focused systematic review of the available trial evidence? This could be information on recent and informal unpublished evidence, or information from registries and other nationally coordinated clinical audits. Any such information must include sufficient detail to allow a judgement to be made as to the quality of the evidence and to allow potential sources of bias to be determined.

Implementation issues

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

The NHS is required by the Department of Health and the Welsh Assembly Government to provide funding and resources for medicines and treatments that have been recommended by NICE technology appraisal guidance. This provision has to be made within 3 months from the date of publication of the guidance.

If the technology is unlikely to be available in sufficient quantity, or the staff and facilities to fulfil the general nature of the guidance cannot be put in place within 3 months, NICE may advise the Department of Health and the Welsh Assembly Government to vary this direction.

Please note that NICE cannot suggest such a variation on the basis of budgetary constraints alone.

How would possible NICE guidance on this technology affect the delivery of care for patients with this condition? Would NHS staff need extra education and training? Would any additional resources be required (for example, facilities or equipment)?

NICE National Institute for Health and Care Excellence

Clinical expert statement

Obinutuzumab for untreated advanced follicular lymphoma [ID1120]

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	**Professor Andrew Pettitt**
2. Name of organisation	University of Liverpool

3. Job title or position	Professor and Consultant Haematologist
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.)	
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.)	
9. In your view, is there an unmet need for patients and healthcare professionals in this condition?	
What is the expected place of the technology in current practice?	

10. How is the condition currently treated in the NHS?	
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	
• 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.)	
• What impact would the technology have on the current pathway of care?	
11. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	

• How does healthcare resource use differ between the technology and current care?	
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	
12. Do you expect the technology to provide clinically meaningful benefits compared with current care?	
• Do you expect the technology to increase length of life more than current care?	

• Do you expect the technology to increase health-related quality of life more than current care?	
13. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	
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	

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?	
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?	
17. Do you consider the technology to be innovative in its potential to make a significant and substantial	

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?	
• Does the use of the technology address any particular unmet need of the patient population?	
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?	
Sources of evidence	

19. Do the clinical trials on the technology reflect current UK clinical practice?	
• If not, how could the results be extrapolated to the UK setting?	
• What, in your view, are the most important outcomes, and were they measured in the trials?	
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	
20. Are you aware of any relevant evidence that might	

not be found by a systematic review of the trial evidence?	
21. Are you aware of any new evidence for the comparator treatment(s) since the publication of NICE technology appraisal guidance [TAXXX]? [delete if there is no NICE guidance for the comparator(s) and renumber subsequent sections]	
22. How do data on real-world experience compare with the trial data?	
Equality	
23a. Are there any potential equality issues that should be	

taken into account when considering this treatment?	
23b. Consider whether these issues are different from issues with current care and why.	
Topic-specific questions	
24. [To be added by technical team if required, after receiving the company submission. For example, if the company has deviated from the scope (particularly with respect to comparators) – check whether this is appropriate. Ask specific, targeted questions such as "Is comparator X [excluded from company submission] considered to be established clinical practice in	

the NHS for treating [condition Y]?"] **if not delete highlighted rows and renumber below**	
Key messages	
25. In up to 5 bullet points, please summarise the key messages of your statement. • • • • •	

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.

NICE National Institute for Health and Care Excellence

Patient expert statement

1020 - Obinutuzumab for untreated advanced follicular lymphoma

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 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 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 10 pages.

ıbout you	
.Your name	**Geraldine Mason**
. Are you (please tick all that pply):	☒ X a patient with the condition? ☐ a carer of a patient with the condition? ☐ a patient organisation employee or volunteer?

fter submission.)	
. How did you gather the ıformation included in your tatement? (please tick all that pply)	☐ I have personal experience of the condition ☐ I have personal experience of the technology being appraised ☐ I have other relevant personal experience. Please specify what other experience: ☐ I am drawing on others' experiences. Please specify how this information was gathered:
.iving with the condition	
. What is it like to live with the ondition? What do carers xperience when caring for omeone with the condition?	It's OK. Side effects and late side effects of Rituximab have been a problem. I have had to substantially modify my diet. I have hypogammagobulinaemia so my immunity is not good – resulting in lots of colds and on one occasion pneumonia, but I did not have to be admitted to hospital. My immunity may be improving now. Also I realise that staying at home is not good for my mental or physical health – getting out is very important. I am getting back into country walking, and cycling, which I used to do a lot. I used to swim out of doors and hopefully can get back to this.
:urrent treatment of the condition in the NHS	
. What do patients or carers ıink of current treatments and are available on the NHS?	Rituximab was an early targeted treatment and is remarkable in its way.
0. Is there an unmet need for atients with this condition?	There's confusion around what late side effects are – do they come on later or are they just side effects which don't go away? I've spoken with a number of patients who have said that they were never the same again after treatment X.
ıdvantages of the technology	

1. What do patients or carers hink are the advantages of the echnology?	Are we talking about chemotherapy such as R-CVP? You don't die. I could have had a DVT, for example.
Disadvantages of the technology	
2. What do patients or carers hink are the disadvantages of he technology?	See above – permanent changes.
Patient population	
3. Are there any groups of atients who might benefit nore or less from the echnology than others? If so, lease describe them and xplain why.	
Equality	
4. Are there any potential quality issues that should be aken into account when	

onsidering this condition and ıe technology?	
)ther issues	
5. Are there any other issues ıat you would like the ommittee to consider?	A similar issue. As with treatments for schizophrenia, the focus is often on addressing symptoms, but not so much attention is paid in research to side effects, as far as I know. But unlike in mental health, this is addressed in discussion with the clinical team, which I think has been very good. At least one of the consultants is happy to spend time working out exactly what the problem is and addressing it.
Topic-specific questions	
6. [To be added by technical eam if required, after receiving ıe company submission. For xample, if the company has eviated from the scope particularly with respect to omparators) – check whether ıis is appropriate. Ask pecific, targeted questions uch as "Is comparator X excluded from company ubmission] considered to be	

stablished clinical practice in ıe NHS for treating [condition]?"] : **not delete highlighted** **ows and renumber below**	

Key messages

7. In up to 5 bullet points, please summarise the key messages of your statement:

- Make sure the patient can have proper unhurried discussions with an interested consultant.
- Make sure patient is working with a Clinical Nurse Specialist with whom they can relate.
- Research on late side effects is needed.
-
-

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.

in collaboration with:

ERASMUS UNIVERSITEIT ROTTERDAM
INSTITUTE OF HEALTH POLICY & MANAGEMENT

Maastricht University

Obinutuzumab for untreated advanced follicular lymphoma

Produced by

Kleijnen Systematic Reviews Ltd. in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University

Authors

Rob Riemsma, Reviews Manager, KSR Ltd, UK
Isaac Corro Ramos, Health Economics Researcher, EUR, NL
Frederick Thielen, Health Economics Researcher, EUR, NL
Debra Fayter, Systematic Reviewer, KSR Ltd, UK
Nigel Armstrong, Health Economist, KSR Ltd, UK
Ching-Yun Wei, Health Economist, KSR Ltd, UK
Ciara Keenan, Systematic Reviewer, KSR Ltd, UK
Vanesa Huertas Carrera, Systematic Reviewer, KSR Ltd, UK
Nasuh Büyükkaramikli, Health Economics Researcher, EUR, NL
Kate Misso, Information Specialist, KSR Ltd, UK
Gill Worthy, Statistician, KSR Ltd, UK
Jos Kleijnen, Director, KSR Ltd, Professor of Systematic Reviews in Health Care, Maastricht University

Correspondence to

Rob Riemsma, Kleijnen Systematic Reviews
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, UK
YO19 6FD

Date completed

18/07/2017

Source of funding: This report was commissioned by the NIHR HTA Programme as project number STA 16/112/02.

Declared competing interests of the authors

None.

Acknowledgements

None.

Commercial in confidence (CiC) data are highlighted in blue throughout the report.

Academic in confidence (AiC) data are highlighted in yellow throughout 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:

Riemsma R, Corro Ramos I, Thielen F, Fayter D, Armstrong N, Wei C-Y, Keenan C, Huertas Carrera V, Büyükkaramikli N, Misso K, Worthy G, Kleijnen J. Obinutuzumab for untreated advanced follicular lymphoma: a Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd, 2017.

Contributions of authors

Rob Riemsma acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Isaac Corro Ramos acted as health economic project lead, contributed to the writing of the report and supervised the health economic part of the project. Nasuh Büyükkaramikli, Frederick Thielen, Ching-Yun Wei and Nigel Armstrong acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Debra Fayter, Ciara Keenan, and Vanesa Huertas Carrera acted as systematic reviewers, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Gill Worthy acted as statistician, critiqued the analyses in the company's submission and contributed to the writing of the report. Kate Misso critiqued the search methods in the submission and contributed to the writing of the report. Jos Kleijnen critiqued the company's definition of the decision problem and their description of the underlying health problem and current service provision, contributed to the writing of the report and supervised the project.

Abbreviations

ADCC	Antibody-dependent cellular cytotoxicity
ADCP	Antibody-dependent cellular phagocytosis
AE	Adverse Events
AESI	Adverse events of special interest
AIC	Akaike information criterion
ALT	Alanine transaminase
ANC	Absolute neutrophil count
ASCO	American Society of Clinical Oncology
ASCT	Autologous stem cell transplantation
ASH	American Society of Hematology
AST	Aspartate transaminase
BCSH	British Committee for Standards in Haematology
BI	Budget impact
BIC	Bayesian information criterion
BM	Bone marrow
BMI	Body mass index
BNF	British National Formulary
BOR	Best overall response
BSA	Body surface area
CADTH	Canadian Agency for Drugs and Technologies in Health
CDC	Complement dependent cytotoxicity
CDF	Cancer Drugs Fund
CE	Cost effectiveness
CEA	Cost effectiveness analysis
CEAC	Cost effectiveness acceptability curve
CHMP	Committee for Medicinal Products for Human Use
CHOP	Cyclophosphamide, doxorubicin, vincristine and prednisone
CHVP	Cyclophosphamide, etoposide, doxorubicin and prednisolone
CI	Confidence Interval
CLL	Chronic lymphocytic leukaemia
CMH	Cochran–Mantel–Haenszel test
COMP	Committee for Orphan Medicinal Products
CR	Complete response
CRD	Centre for Reviews and Dissemination
CrI	Credible interval
CS	Company's submission
CSR	Clinical study report
CT	Computed tomography
CVP	Cyclophosphamide, vincristine, and prednisone
DFS	Disease-free survival
DLBCL	Diffuse large B cell lymphoma
DSU	Decision Support Unit
ECOG	Eastern Cooperative Oncology Group
EFS	Event-free survival
EHA	European Haematology Association
EMA	European Medicines Agency
EOI	End of induction
EOMR	End of maintenance response
EPAR	European public assessment report
EQ-5D	European Quality of Life-5 Dimensions
ERG	Evidence Review Group
ESMO	European Society for Medical Oncology
EUR	Erasmus University Rotterdam

FACT-G	Functional Assessment of Cancer Therapy - General
FACT-Lym	Functional Assessment of Cancer Therapy for Patients with Lymphoma
FACT-TOI	Functional Assessment of Cancer Therapy - Trial Outcome Index
FC	Fludarabine and cyclophosphamide
FDA	Food and Drug Administration
FDG-PET	Fludeoxyglucose positron-emission tomography
FL	Follicular lymphoma
FLIPI	Follicular Lymphoma International Predictive Index
FU	Follow up
G-CSF	Granulocyte colony stimulating factor
GI	Gastrointestinal
HDT	High-dose therapy
HMRN	Haematological Malignancy Research Network
HR	Hazard ratio
HRQoL	Health Related Quality of Life
HSUV	Health state utility
HTA	Health Technology Assessment
IC	Indirect Comparison
ICER	Incremental Cost Effectiveness Ratio
ICML	International Conference on Malignant Lymphoma
ICTRP	International Clinical Trials Registry Platform
IDMC	Independent Data Monitoring Committee
INAHTA	International Network of Agencies for Health Technology Assessment
iNHL	Indolent non-Hodgkin lymphoma
INR	International normalised ratio
IOG	Improving Outcomes Guidance
IRC	Independent review committee
IRR	Infusion related reactions
ISPOR	International Society For Pharmacoeconomics and Outcomes Research
ITT	Intention to Treat
IVRS	Interactive Voice Response System
KM	Kaplan–Meier
KSR	Kleijnen Systematic Reviews
LAA	Last antibody administration
LDH	Lactate dehydrogenase
LVEF	Left ventricular ejection fraction
LYG	Life years gained
MALT	Mucosa associated lymphoma tissue
MCP	Mitoxantrone, chlorambucil and prednisolone
MeSH	Medical Subject Headings
MHRA	Medicines and Healthcare Products Regulatory Agency
MRD	Minimal residual disease
MRI	Magnetic Resonance Imaging
MTC	Mixed Treatment Comparison
MTX	Methotrexate
MUGA	Multigated radionuclide angiography
MZL	Marginal zone lymphoma
NA	Not applicable
NALT	New anti-lymphoma treatment
NCCN	National Comprehensive Cancer Network
NE	Not estimated
NIH	National Institute of Health (US)
NHS	National Health Services
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research

NR	Not Reported
Obin	Obinutuzumab (maintenance monotherapy)
Obin-chemo	Obinutuzumab in combination with chemotherapy as induction therapy
ORR	Overall response rate
OS	Overall survival
PAS	Patient access scheme
PB	Peripheral Blood
PCR	Polymerase chain reaction
PD	Progressed disease
PET	Positron-emission tomography
PFS	Progression-free survival
PML	Progressive multifocal leukoencephalopathy
PPS	Post-progression survival
PR	Partial response
PRESS	Peer Review of Electronic Search Strategies
PRO	Patient-reported outcome
PSA	Probabilistic Sensitivity Analyses
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
PTT	Partial thromboplastin time
QALY(s)	Quality-adjusted Life Year(s)
QoL	Quality of life
RCT	Randomised Controlled Trial
R	Rituximab (as maintenance monotherapy in GALLIUM)
R-chemo	Rituximab in combination with chemotherapy as induction therapy
RR	Relative Risk; Risk Ratio
SACT	Systemic Anti-Cancer Therapy (chemotherapy dataset)
SAE	Serious Adverse Events
SC	Sub-cutaneous
SD	Stable disease
SE	Standard error
SLL	Small lymphocytic lymphoma
SOC	Standard of care
SPC	Summary of product characteristics
STA	Single Technology Appraisal
TEAEs	Treatment-emergent adverse events
TLS	Tumour lysis syndrome
TTTD	Time-to-treatment-discontinuation
UK	United Kingdom
ULN	Upper limit of normal
UMC	University Medical Centre
WHO	World Health Organisation
WM	Waldenstrom macroglobulinaemia

Table of Contents

Table of Tables

Table of Figures

1. SUMMARY

1.1 Critique of the decision problem in the company's submission

The NICE scope describes the decision problem as obinutuzumab in combination with chemotherapy, with or without obinutuzumab maintenance therapy for people with untreated advanced follicular lymphoma. The comparators are described as rituximab monotherapy, rituximab-based chemotherapy, with or without rituximab maintenance treatment and bendamustine monotherapy.

The anticipated licence for obinutuzumab is in combination with chemotherapy (CHOP, CVP or bendamustine), followed by maintenance therapy in patients achieving a response, for the treatment of patients with previously untreated advanced follicular lymphoma except for follicular lymphoma (FL) grade 3b.

The final wording at CHMP opinion, received from the company after completion of this report as part of the check for factual inaccuracies, is as follows: "Gazyvaro in combination with chemotherapy, followed by Gazyvaro maintenance therapy in patients achieving a response, is indicated for the treatment of patients with previously untreated advanced follicular lymphoma."

The anticipated licence and the main trial in the submission, GALLIUM, excludes patients with follicular lymphoma (FL) grade 3b. Apart from this, the population in the submission matches the scope. The anticipated licence and the included trial only include obinutuzumab induction followed by maintenance therapy for responders. The company does not provide any separate evidence for obinutuzumab induction therapy without maintenance therapy.

The company has presented evidence for one of the comparators in the scope: rituximab-based chemotherapy followed by rituximab maintenance treatment (this is the comparator in the GALLIUM trial). No evidence has been presented for rituximab mono-therapy and bendamustine mono-therapy.

The NICE final scope lists the following outcome measures: overall survival, progression-free survival, overall response rate, adverse effects of treatment and health-related quality of life. These outcomes are reported in the company submission (CS). However, overall survival (OS) data of the GALLIUM trial are still immature at the cut-off date for primary analysis (January 2016), with fewer than 20% of patients followed for survival for more than four years.

1.2 Summary of clinical effectiveness evidence submitted by the company

One phase III, open-label randomised controlled trial (RCT), GALLIUM with 1,202 previously untreated adult participants with follicular lymphoma was presented as the main source of evidence in the submission.

Patients were randomly assigned in a 1:1 ratio to either obinutuzumab + chemotherapy followed by obinutuzumab monotherapy maintenance in responders, or to rituximab + chemotherapy followed by rituximab monotherapy maintenance in responders. Stratification factors for randomisation were: chemotherapy regimen, Follicular Lymphoma International Predictive Index (FLIPI) (low or high), and geographic region (Western Europe, Eastern Europe, South and Central America, North America, other). Prior to the initiation of the study, each site chose one of three chemotherapy regimens (CHOP, CVP, or bendamustine) that was considered to be the standard of care for follicular lymphoma; all patients with follicular lymphoma at that site received the chosen chemotherapy regimen for the duration of the study.

In the obin-chemo arm, eight to 10 doses of obinutuzumab at 1,000 mg were administered by IV infusion with the accompanying chemotherapy regimen during induction. Patients randomised to receive obin-chemo who achieved a complete response (CR) or partial response (PR) at the end of induction therapy continued to receive obin-maintenance at 1,000 mg every two months until disease progression, or for two years. In the r-chemo arm, six to eight doses of R at 375 mg/m2 were administered by IV infusion with the accompanying chemotherapy regimen during induction. Patients randomised to receive r-chemo who achieved a CR or PR at the end of induction therapy continued to receive r-maintenance at 375 mg/m2 every two months until disease progression, or for two years.

The primary outcome was investigator-assessed progression-free survival (PFS). Key secondary outcomes included PFS assessed by independent review committee (IRC), overall survival and response rates. Health-related quality of life was also assessed using a disease-specific tool (FACT-Lym) and EQ-5D. The submission focused on results of effectiveness on a data cut of January 2016. On request, the company provided full results for the later cut-off of September 2016. We have provided results for both time points in this report.

Overall, obinutuzumab was superior to rituximab for PFS (HR = 0.72 (0.56 to 0.93)) for the latest cut-off using IRC data. Although outcomes relating to progression were positive, no differential effects on HRQoL between groups were identified. Overall survival data in GALLIUM were not mature. At the updated clinical cut-off date (10 September 2016), 95 patients (7.9% of the FL population) had died. Although overall rates of adverse events between groups were similar, a higher rate of serious adverse events was noted with obinutuzumab (46.1% vs 39.9%). These led to a higher rate of dose withdrawal, reduction or interruption in the obinutuzumab group. Grade 3 to 5 adverse events (AEs) were also more frequent with obinutuzumab (74.6% vs 67.8%).

The GALLIUM study is ongoing. At the time of the clinical cut-off date (31 January 2016), 114 patients with FL were still undergoing maintenance treatment (54 in the r-chemo arm and 60 in the o-chemo arm).

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

The original literature searches reported in the CS contained several typographical mistakes and consequential errors in database command language/syntax which the ERG raised during the clarification process. The company re-structured and re-executed almost all of the searches, and provided replacement strategies in the clarification response. Consequential problems with wildcard use remained in some strategies. A range of databases were searched, and additional searches of conference proceedings, trials registers and websites were conducted. Searches were carried out in accordance with the NICE guide to the methods of technology appraisal sections 5.2.2 and 5.2.4.

The CS and response to clarification provided sufficient details for the ERG to critique the clinical effectiveness.

The clinical effectiveness evidence in the submission is based on one trial, GALLIUM. Although GALLIUM is a good quality RCT, a number of limitations were identified by the ERG. The trial was open-label, therefore results based on independent review will be less prone to bias than results based on investigator assessment. In the trial, obinutuzumab and its comparator rituximab could be given to patients with three different chemotherapy regimens (CHOP, CVP and bendamustine). In the trial approximately 57% received bendamustine, 33% CHOP and 10% CVP. The breakdown of the chemotherapy used may not be reflective of the UK. The trial was not designed to investigate differences in chemotherapy regimens so any variation in results between chemotherapy regimens may reflect genuine differences of effectiveness or patient selection factors.

Although GALLIUM had a reasonable follow-up duration, data were not fully mature for the main outcomes. Median progression-free survival (PFS) could not be determined. Although outcomes relating to progression were positive, no differential effects on HRQoL were identified. The committee will need to consider the possible relationship between improvements in PFS and subsequent improvements in overall survival as overall survival data in GALLIUM were not mature. GALLIUM is an ongoing trial which should provide further, more mature results. Finally, the higher rate of serious and higher grade adverse events with obinutuzumab needs to be considered in terms of management of the disease and acceptability to patients.

1.4 Summary of cost effectiveness evidence submitted by the company

The company developed a de novo economic model to assess the cost effectiveness of obinutuzumab in combination with chemotherapy (CVP, CHOP or bendamustine), followed by obinutuzumab maintenance therapy in patients achieving a response, compared to rituximab-based chemotherapy, with rituximab maintenance treatment, in people with untreated advanced follicular lymphoma except for FL grade 3b (i.e. the population in the GALLIUM trial).

The model developed for this submission was a four-state cohort transition (Markov) model with monthly cycles and a time horizon of 40 years. The health states considered in the model are progression free (PFS) (on/off treatment), two progressed disease (PD) states, early PD and late PD and death. All patients begin in the PFS health state on treatment and are assigned to a PFS 'on-treatment' utility value and treatment costs while on therapy. Time to treatment discontinuation is based on the actual observation from the GALLIUM study for both arms. Specifically, as per license indication, only patients responding to induction received maintenance. Maintenance was only offered until progression or for a maximum of two years; then it is said that treatment is completed. When patients complete or discontinue treatment in the PFS state, they are considered off treatment and assigned an 'off treatment' PFS utility value and costs for ongoing monitoring in supportive care. Patients can either remain in PFS (on- or off-treatment) or exit the state due to disease progression or death. Two progressed disease states were introduced to account for different outcomes and costs to the cohorts of patients who experience an early or a late progression. Once patients enter any of the two PD states, patients can only remain in their corresponding PD state until death. The model also includes the most common adverse events observed in the GALLIUM trial. The cost and disutility effects (the latter only in scenario analysis) of these adverse events were incorporated in the PFS (on-treatment) health state for a maximum of two years.

Clinical parameters for the model were derived from the GALLIUM trial data when these were considered mature enough to provide robust estimates. Thus, GALLIUM data were used to estimate time to treatment discontinuation (TTTD), PFS and post progression survival (PPS) for early progressed disease. The investigator (INV) assessed PFS data (PFS-INV) were used, corresponding to the primary endpoint. The extrapolation beyond the observed period in the GALLIUM trial was based on parametric functions. The latest available data cut of GALLIUM with a clinical cut-off date of 10 September 2016 was used. External data were used to populate PPS for late progressed disease using long-term data from the PRIMA trial.

The model uses EQ-5D utilities collected from the GALLIUM trial for the PFS health state. Since long-term EQ-5D utility scores collected in the GALLIUM trial were considered immature by the company, post progression utilities were sourced from the literature.

Health state related costs consisted of medication costs (induction and maintenance), supportive care costs, subsequent treatment costs in PD, and adverse event costs. Relevant medication costs included

costs of obinutuzumab, bendamustine, CHOP, CVP, and rituximab. Resource use was derived from UK reference costs.

The results of the base-case cost effectiveness analysis showed that obin-chemo+obin resulted in a total cost of [redacted] and 10.01 QALYs. The comparator, R-chemo+R, resulted in a total cost of [redacted] and 9.23 QALYs. Thus, obin-chemo+obin produced 0.78 additional QALYs at an incremental cost of [redacted] when compared to R-chemo+R, leading to an ICER of [redacted].

The probabilistic sensitivity analysis (PSA) results presented by the company estimated that the probability that obin-chemo+obin is cost effective compared to R-chemo+R is approximately [redacted] at a threshold of £30,000 per QALY gained. The result of the deterministic sensitivity and scenario analyses showed that the incremental cost effectiveness ratios (ICERs) remained below [redacted] and were close to the base-case value in most cases. The most influential parameter was the duration of the treatment effect, whose variation resulted in a wide range of possible values of the ICER.

The company conducted several scenario analyses to explore the impact on the cost effectiveness results of several of the structural uncertainties which are present in the economic evaluation. The company considered two scenarios in the CS. The main purpose of the first scenario was to compare the company de novo model with the latest economic analysis of rituximab in combination with chemotherapy in the first-line treatment of FL. In this scenario, the ICER decreased to [redacted]. In the second scenario, the company assumed no difference in costs and QALYs gained post progression. The resulting ICER was [redacted], thus lower than the ICER in the base-case scenario. Additionally, the company ran three scenarios as requested by the ERG in the clarification letter. In the first additional scenario the age at baseline was increased by 4.7 years in line with the data reported in the Haematological Malignancy Research Network (HMRN) database. In this scenario, the ICER increased to [redacted]. In the next scenario, the cost for chemotherapy, administration and adverse events were re-weighted according a different chemotherapy distribution (based on UK market research). This resulted in a smaller cost difference between the R-chemo+R and the obin-chemo+obin arm in drug acquisition, administration and AE costs, thus decreasing the ICER to [redacted]. In the final scenario, the company modelled PFS and PPS mortality separately per treatment arm, which resulted in an increased ICER ([redacted]).

1.5 Summary of the ERG's critique of cost effectiveness evidence submitted

The ERG's major concern with respect to the company submission was the validity of some assumptions regarding the implementation of the treatment effectiveness in the economic model. These were related with the duration of the treatment effect and the choice of PFS data (local investigator or independent review committee).

The assumption of finite duration of treatment effect on PFS is the main driver of the cost effectiveness results. In the absence of long-term data in the GALLIUM trial, this assumption was made based on the PRIMA trial. However, it should be noted that the PRIMA trial compares rituximab maintenance after induction chemotherapy with observation (i.e. no maintenance), while in the GALLIUM trial maintenance with obinutuzumab and rituximab are considered. Whether the same long-term treatment effect applies to obinutuzumab compared to rituximab is therefore speculative. While the company presented this choice as conservative, given that the clinical advisors and the literature consulted suggested that there is no evidence of a finite duration of treatment effect in treatments of FL, it should be noted that there is no evidence of the opposite either (possibly due to the limited long-term follow-up data). According to the evidence in the company submission, the ERG could not propose an alternative estimate for the treatment effect duration that could have been considered robust. For that reason, the ERG explored the impact of this parameter in a threshold analysis. Based on its results, the

ERG considers that assuming a treatment effect for five years (which also coincides with the longest follow up in the GALLIUM trial) could have been seen as a more conservative approach than the one presented in the company's base-case.

The investigator (INV) assessed PFS data (PFS-INV) was used in the company's base-case analysis. The ERG considers that independent review committee (IRC) assessed PFS (PFS-IRC) data should have been used for the company's base-case analysis because the GALLIUM trial was open-label and the results based on independent review are less prone to bias than investigator results. The use of PFS-IRC data was investigated in a scenario analysis by the company. In this scenario, the company assumed the same parametric distribution for PFS as in the company's base case: the exponential. However, the ERG considers that this is not correct. Since the PFS-IRC dataset is different from the PFS-INV dataset, the goodness of fit for the PFS-IRC data should have been reassessed.

Other concerns of the ERG were related to the generalisability to UK clinical practice (in particular the baseline age of the patient population and the proportion of patients per chemotherapy method) and the estimation of utility and cost input parameters.

The proportion of UK patients in the GALLIUM trial (21%) seems reasonable and nearly half of the patients are from Western Europe. However, the company acknowledged that the GALLIUM cohort might be younger than the average UK patient. This was also confirmed by some of the clinical experts consulted by the company.

The proportion of patients treated with each chemotherapy method (bendamustine, CHOP, CVP) in the GALLIUM trial and the proportion reported for the general UK population were quite different and might indicate that the proportions used in the GALLIUM trial are not reflective of the UK clinical practice. In the clarification response, the company mentioned that, according to the discussions in the advisory board, there are local variations in clinical practice with respect to chemotherapy use and therefore, the appropriate representative average use of the three chemotherapy regimens has some uncertainty. This implies that whether the proportions used in the GALLIUM trial are reflective or not of the UK clinical practice is also uncertain. Since GALLIUM was not powered to detect differences between the three chemotherapy methods and patients were not randomised to chemotherapies, the ERG considers that it is not feasible to conduct a robust scenario analysis where PFS and OS estimates are obtained with a different proportion of chemotherapy regimens. The only feasible scenario analysis may be then to assume equal clinical outcomes while considering chemotherapy, administration and AE costs according to an alternative distribution of patients per chemotherapy group. However, if there is any treatment effect due to the underlying chemotherapy method, this would not be possible to detect with the current analyses.

There were several concerns regarding the assumptions made for the utilities in the company's model. These were related with the applicability (or not) of an UK tariff for non-UK utility values, reliability of the utilities for the progressed disease health state and adjusting utilities for a decline in age.

It seems that all available data from GALLIUM were used regardless of the geographical region. It is not clear whether UK tariff has been applied to the GALLIUM utilities or not. This is not mentioned in the company submission and the utilities reported by the company could not be verified by the ERG. In the updated CSR document provided by the company, the ERG could not find any UK-specific EQ-5D data. There were nevertheless several tables reporting EQ-5D values for Western Europe and it was observed that these seem to be lower than the overall ones. The reasons for the differences between utilities of Western European patients and others were not clear to the ERG.

In spite of being unpublished, inconsistent with the results of the GALLIUM trial and unverifiable (by the ERG), the company relied on the utility values reported by Wild et al. 2006. The ERG judged the derivation and choice of EQ-5D utility values for the PD health-state in the CS as non-transparent and non-replicable. However, given the available evidence, the ERG was not able to decide which of the estimates reported in the literature were the most reliable and representative for the patient population. For that reason, the ERG used the values from Wild et al. to model utilities in the PD health-state in its preferred base-case and explored alternative options in scenario analyses.

The ERG does not agree with the company's assumption of not adjusting the utility values for a decline in age. Seeing the age distribution in the GALLIUM trial, it seems very unlikely that the trial was powered to detect differences in utilities for different age groups. Therefore, any assumption based on this does not seem to be valid. For that reason, the ERG considers that the decline in age for utilities should have been included in the base-case analysis, which would also result in a more conservative ICER.

The ERG considers that, in general, the company provided a solid overview on the costs and resource use used in the economic model. The ERG verified the references of all available sources and corrected some inconsistencies in the calculation of some cost items.

1.6 ERG commentary on the robustness of evidence submitted by the company

1.6.1 Strengths

Searches were carried out in line with the NICE guide to the methods of technology appraisal sections 5.2.2 and 5.2.4. The company's clarification response provided sufficient details for the ERG to appraise the searches. Additional searches were carried out for conference abstracts.

The clinical evidence is based on a good quality randomised controlled trial including 1,202 patients with follicular lymphoma. The comparator arm was rituximab, a valid comparator for this appraisal and in clinical practice. Outcomes assessed reflect the scope and are relevant to patients in practice.

The cost effectiveness section of the company submission is well structured and the cost effectiveness analyses have been reported transparently. Furthermore, the analyses of the survival data were correctly performed, following the guidance from the NICE Decision Support Unit. For the extrapolation of progression free survival beyond the trial period, parametric functions were fitted simultaneously for both treatment arms data, with treatment as a covariate in the model, which allowed accommodating both proportional hazards and accelerated failure time models. Additionally, the structure of the model developed by the company is in line with other, commonly used, Markov models for progression in oncology but it has the advantage of incorporating early and late progressed disease health states, which seems to be appropriate for the decision problem at hand. The model also includes relevant adverse events, utilities and costs. Sensitivity analyses were performed on the model parameters and the results were robust to most of the structural assumptions.

1.6.2 Weaknesses and areas of uncertainty

The Resource Use searches and all Cochrane Library/CENTRAL/NHS EED search strategies contained errors in wildcard use and truncation, which may have affected strategy recall. The clinical effectiveness, cost effectiveness and resource use search include a Line of Treatment facet that the ERG felt was overly restrictive. Searches for adverse events data, non-randomised and non-controlled evidence, and indirect and mixed treatment comparisons were not conducted. It is possible that relevant evidence may have been missed as a consequence of this.

Although GALLIUM had a reasonable follow up duration, data were not fully mature for the main outcomes. Median progression-free survival (PFS) could not be determined and overall survival data in GALLIUM were not mature.

GALLIUM is an ongoing trial which should provide, further, more mature results. Further research also might include an investigation comparing obinutuzumab with different chemotherapy regimens (CHOP, CVP and bendamustine).

The main weakness of the cost effectiveness section of the company submission is the reliance on assumptions that could not be verified with the presented evidence. In particular, the duration of the treatment effect and the choice of the progression free survival probability distribution might have a major impact on the cost effectiveness results.

The health-related quality of life section of the company submission is sometimes lacking transparency. It remained unclear whether a UK tariff was applied or not and whether the effects of adverse events while on treatment were captured in the utility values collected in the GALLIUM trial. Furthermore, despite being unpublished, inconsistent with the results of the GALLIUM trial and unverifiable (by the ERG), the company relied on the utility values reported by Wild et al. to inform the utilities assigned to the progressed disease health state. The choice of these utilities might also have a significant impact on the cost effectiveness results.

Finally, it also remained uncertain whether the proportions of patients treated with each chemotherapy option (bendamustine, CHOP, CVP) used in the GALLIUM trial are reflective of the UK clinical practice or not. Since GALLIUM was not powered to detect treatment effect differences for each of the three chemotherapy options and patients were not randomised to/within the chemotherapies, it was not feasible to conduct a scenario analysis where the PFS and OS estimates are based on a different proportion of chemotherapy regimens (e.g. inspired from UK clinical practice). However, if there was any treatment effect due to the underlying chemotherapy method, this would not be possible to detect with the current analyses.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG preferred base-case resulted in an ICER of [redacted] per QALY gained. The ERG's most influential adjustments were 1) choosing PFS-IRC data and a Weibull distribution for PFS extrapolation; 2) applying a utility decrement by age; 3) increasing age at baseline and 4) considering different mortality rates per treatment arm. From the PSA results, the probability that obin-chemo+obin is cost effective compared to R-chemo+R was approximately [redacted] at a £30,000 per QALY gained threshold. Thus, an absolute reduction of [redacted] in the cost effectiveness probability compared to the company base-case probability at the same threshold. The key findings from company and ERG preferred base-case analyses are shown in Table 1.1.

Table 1.1: Key finding from company and ERG analyses

Scenarios	Obin-chemo+obin		R-chemo+R		Inc Costs (£)	Inc QALYs	ICER (£)
	Total Costs (£)	Total QALYs	Total Costs (£)	Total QALYs			
CS base-case	[redacted]	10.01	[redacted]	9.23	[redacted]	0.78	[redacted]
ERG preferred base-case	[redacted]	9.12	[redacted]	8.58	[redacted]	0.53	[redacted]

CS = company submission; ERG = evidence review group; ICER = incremental cost effectiveness ratio; QALYs = quality adjusted life years.

The ERG conducted several scenario analyses where the structural uncertainties in the ERG preferred base-case were explored. These analyses were categorized in four groups: clinical effectiveness, utilities, demographic characteristics and costs.

The ERG first performed a threshold analysis on the duration of the treatment effect. In particular, assuming a duration of the treatment effect of two years (the maximum time on maintenance assumed in the GALLIUM trial) resulted in an ICER of [redacted]. Assuming five years, which is the longest follow-up in the GALLIUM trial (and presented as scenario analysis in the company submission), resulted in an ICER of [redacted]. Moreover, a treatment effect duration of [redacted] was the maximum value assumed where the ICER was above the £30,000 threshold. Additionally, the ERG performed a scenario where the duration of the treatment effect was included in the PSA. This was modelled as a uniform distribution between 0 and 18 years. This scenario resulted in a probabilistic ICER ([redacted]) which was [redacted] higher than the ERG preferred base-case probabilistic ICER ([redacted]). The uncertainty associated to this scenario was increased when compared to the ERG base-case, and at the £30,000 threshold ICER the probability that obin-chemo+obin is cost effective was approximately [redacted]. Thus, an absolute reduction of [redacted] and [redacted] in the cost effectiveness probability compared to the ERG and company base-case probability at the same threshold, respectively. In the scenario where a Gompertz distribution was chosen to model progression free survival the ICER was [redacted]. Assuming PFS-INV data to model PFS and a Weibull distribution to extrapolate resulted in an ICER of [redacted]. Finally, when a pooled mortality for both treatment arms was assumed, the ICER obtained was [redacted].

Within the set of scenarios performed on utilities, the ICER ranged from [redacted], when the utilities collected in the GALLIUM trial were used for the PFS health state and the utilities reported in Bec et al. were used for the PD health state, to [redacted] when GALLIUM utilities were used for both PFS and PD health states. This showed that the ICER is sensitive to changes in utilities. In general, it was observed that assuming higher utility values for the PD health state resulted in a higher ICER.

In another scenario, the baseline age was decreased to the one observed in the GALLIUM trial. The resulting ICER was [redacted]. It seems that the ICER from the ERG preferred base-case is less sensitive to changes in baseline age than the ICER from the company's base-case. This may be explained by the inclusion of utility decrement with age in the ERG preferred scenario.

From the results of the cost-related scenarios, it seems that the ICER might be sensitive to changes in the distribution of patients per chemotherapy group, with CHOP and bendamustine more expensive options and CVP the least expensive. In the hypothetical situation where all patients were assigned to

just one chemotherapy group, the ICERs obtained were [redacted] for bendamustine, [redacted] for CHOP and [redacted] for CVP.

The ERG has drawn attention to a number of parameters for which it is believed that there is uncertainty and therefore, they should have been included in the probabilistic analyses. Most of these parameters could not be included due to lack of data and time constraints. While this is expected to have a minor/moderate impact on the ICER, the current probabilistic results are likely to underestimate the uncertainty around the model results. However, the ERG considers it difficult to ascertain to what extent this is underestimated.

In conclusion, the ERG base-case analysis resulted in an ICER of [redacted] per QALY gained. This ICER value is lower than the £30,000 per QALY threshold. Although the ICER seems to be robust to most of the structural changes explored by the ERG, it is possible that different choices for the treatment effect duration and the utilities for the PD health state would result in an ICER above the £30,000 per QALY threshold.

2. BACKGROUND

In this report the ERG provides a review of the evidence submitted by Roche in support of obinutuzumab, trade name Gazyvaro®, a Type II anti-CD20 antibody for the treatment of patients with advanced follicular lymphoma. In this section we outline and critique the company's description of the underlying health problem and the overview of current service provision. The information is taken from Chapter 3 of the company submission[1] (CS) with sections referenced as appropriate.

2.1 Critique of company's description of underlying health problem

The underlying health problem of this appraisal is follicular lymphoma (FL), a subtype of indolent non-Hodgkin lymphoma (iNHL).

According to the CS *"The Haematological Malignancy Research Network (HMRN) estimate that there will be 1,900 new cases of FL each year in the UK,"* The CS further states that *"In 2015, 2,142 new cases of FL were registered in England (Office for National Statistics, 2017[2]). The 10-year prevalence is estimated at 15,008 cases (25.7 patients per 100,000 people."* The CS also comments that "*The median age at diagnosis of FL in the UK is approximately 65 years old*".[1]

The CS describes the grading of FL (Grades 1 to 3a indolent disease and grade 3b aggressive lymphoma) and the staging of FL (I to IV) according to the Ann-Arbor classification. The CS notes that a patient typically presents with advanced stage disease with multiple sites of lymph adenopathy and / or bone marrow disease. "*At diagnosis, the majority of people with FL have advanced (stage III-IV Ann Arbor stage disease); bone marrow involvement is also common and present in more than 50% of patients."*[1]

It is noted that *"Patients with advanced stage FL are usually considered incurable with standard therapeutic approaches therefore treatment generally attempts to control the disease. FL is typified by a chronic course comprising of repeated relapses, treatment and progression."*[1]

Section 3.2 of the CS describes the effects of FL on patients, carers and society. Several factors impacting on quality of life in iNHL are discussed including unpredictable relapses and associated symptoms, repeated courses of treatment and toxicity of treatment. Furthermore the CS suggested a high level of dependency in patients with iNHL. "*A cross-sectional survey of iNHL patients identified that almost one-quarter of patients depended on caregiver assistance, with the majority (74%) being unpaid care provided by a spouse, partner, relative or friend*."[1]

The CS notes that "*The progression of FL varies among patients depending on the speed of tumour growth and involvement of other organs. Approximately 20% of FL patients who receive immunochemotherapy still suffer PD within two years from diagnosis (Casulo et al., 2015b[3])*"

The CS highlights that a small percentage of patients with FL will transform to more aggressive forms such as diffuse large B cell lymphoma (DLBCL) and therefore have a worse outcome. "*Recent studies report a risk of transformation of about 2% to 3% per year through at least 10 to 15 years of diagnosis*"

The CS describes the prognosis for survival in FL *"Generally, median life expectancy ranges have been reported from 8–12 years after diagnosis, although this has extended to around 15 years in the post-rituximab era.[4]"* "*The HMRN estimate the 5-year survival rate of patients with FL in the UK to be 87.2% (Haematological Malignancy Research Network 2017d[5]).*"

The CS mentions the strategies to predict survival i.e. the Follicular Lymphoma International Predictive Index (FLIPI) and the revised FLIPI2 and outlines the differences between them. The role of minimal residual disease (MRD) in predicting prognosis is also discussed.

The CS cites data from the US National LymphoCare study showing that patients who progress within two years of treatment have a poorer prognosis than responders: "*The US National LymphoCare study (analysing 588 patients with stage 2–4 FL having received first-line R-CHOP) demonstrated that 5-year survival among patients with disease progression within 2 years of treatment was lower compared with those without disease progression, 50% vs 90% respectively"*

ERG comment:

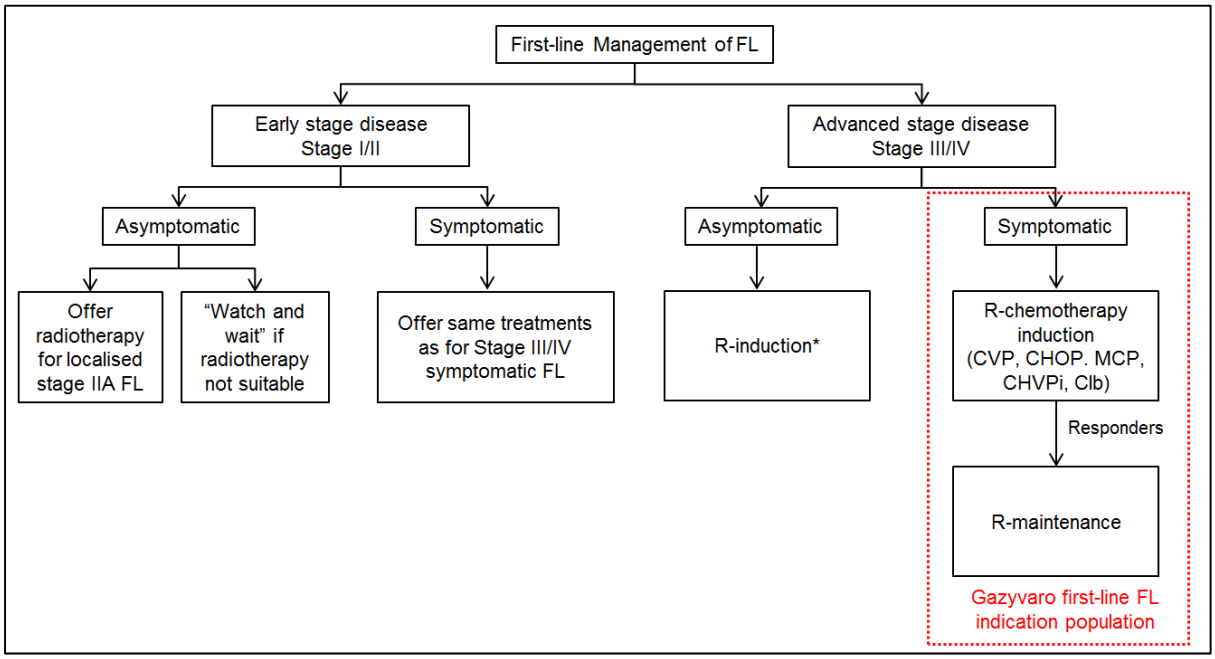
The company provided a good overview of the underlying health problem. The ERG checked the references provided to support the statements in the company submission. In general these were found to be appropriate. However, the ERG noted a small number of points to take into consideration:

- The company presented evidence for the link between progression-free survival and overall survival. The ERG noted that the overall survival rate of 50% with disease progression and 90% without progression in the US National LymphoCare study was before adjustment for FLIPI. However the authors of the study stated that "*This trend was maintained after we adjusted for FL International Prognostic Index (hazard ratio, 6.44; 95% CI, 4.33 to 9.58)*." The relevance of the US LymphoCare study is discussed in section 5 of this report.
- The cross-sectional survey cited by the company as evidence that almost one-quarter of patients with iNHL depend on caregiver assistance was completed by 84 patients of whom 46 (54.8%) had follicular lymphoma.[6] Numbers of patients with FL may be too small in this study to give a reliable percentage dependent on carers.

2.2 Critique of company's overview of current service provision

Figure 2.1 shows the treatment pathway for patients with previously untreated follicular lymphoma, based on NICE technology appraisal guidance 243[7] with the proposed position of obinutuzumab (Gazyvaro). The company submission (CS) describes the intervention as obinutuzumab in combination with chemotherapy, followed by obinutuzumab for maintenance in patients with previously untreated advanced follicular lymphoma, which is in line with the NICE scope.

Figure 2.1: Treatment pathway based on NICE recommendation 243 for patients with previously untreated advanced follicular lymphoma.



Source: Figure 4 of the CS
*Rituximab does not have a UK Marketing Authorisation for this indication
CVP = cyclophosphamide, vincristine and prednisolone; CHOP = cyclophosphamide, doxorubicin, vincristine and prednisolone; NICE = National Institute for Health and Care Excellence

The proposed positioning of the new technology is for advanced stage symptomatic follicular lymphoma which is currently treated by rituximab chemotherapy (R-Chemo), the standard treatment. For responders to initial therapy with R-chemo, maintenance with rituximab may be offered.

The CS also provides an estimated breakdown of the type of chemotherapy used with rituximab at the first line in the UK based on market research and in conjunction with an advisory board of experts. See Table 2.1.

Table 2.1: First-line regimens in UK clinical practice for FL

Regimen	Proportion use, %
Induction	
R-CVP	36
R-CHOP	13
R-bendamustine	29
R-FC	8
R-other	2
FC	11
Other	1

Advisers to the company stated that "*R-CVP and R-bendamustine are the most commonly used induction immunochemotherapy regimens in the UK, with R-CHOP retained for use in patients at high risk of transformation.*"

The CS states that rituximab induction with chemotherapy has been shown to improve progression-free survival and overall survival in trials. Maintenance treatment with rituximab monotherapy has shown

improvements in PFS but benefits in terms of overall survival have not been shown in randomised controlled trials.

The case for the need for obinutuzumab is made. The CS states that "*there is a need for first-line FL treatments that can result in longer remissions and longer time to next lymphoma treatment, and fewer patients requiring treatment in a relapse setting*."

If obinutuzumab were approved by NICE the company note that "*There are no significant changes to the provision of services and patients management. However, Gazyvaro requires additional administration in induction in combination with chemotherapy in comparison to R-chemo. Furthermore, patients in England can be offered the subcutaneous formulation of MabThera for maintenance treatment after response to R-chemo induction, whereas Gazyvaro requires IV administration in maintenance. The respective cost implications were accounted for in the economic analysis*."[1]

ERG comment:

- The case for approval of obinutuzumab is based on superior performance to rituximab as both interventions target the same population of patients with symptomatic FL.
- The ERG asked for clarification on the composition of the advisory board and methodology used in the consultation of clinical experts who were asked to comment on treatments in current practice. The company replied that "*An expert advisory board was consulted at a one-day meeting in April 2017. The panel consisted of …. Consultant haematologists specialising in the management of patients with FL, many of whom have experience of obinutuzumab from clinical trials*."[8] Although the CS provides details surrounding their expertise and employment, the company could have provided more details about how opinions were collected and how many clinicians were involved in each decision.
- A comparison of the breakdown of treatments derived here and in the main trial GALLIUM is discussed in section 4 of this report.
- The ERG draws the attention of the committee to the additional administration costs of obinutuzumab compared to rituximab which are included in the economic model.

3. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 3.1: Statement of the decision problem (as presented by the company)

	Final scope issued by NICE	Decision problem addressed in the company submission	Companies rationale if different from the final NICE scope	ERG comment
Population (s)	People with untreated advanced follicular lymphoma	People with untreated advanced follicular lymphoma	No difference	In line with the scope of the decision problem.
Intervention	Obinutuzumab in combination with chemotherapy, with or without obinutuzumab maintenance therapy	Gazyvaro (obinutuzumab) in combination with chemotherapy (CVP, CHOP or bendamustine), followed by Gazyvaro maintenance therapy in patients achieving a response	Align with wording of expected Marketing Authorisation	Chemotherapy has been specified by the company as CVP, CHOP or bendamustine. No evidence has been presented for obinutuzumab in combination with chemotherapy, without obinutuzumab maintenance therapy. This is in line with the wording of expected Marketing Authorisation
Comparator (s)	• Rituximab monotherapy (does not currently have a Marketing Authorisation in the UK for this indication) • Rituximab-based chemotherapy, with or without rituximab maintenance treatment • Bendamustine monotherapy (does not currently have a Marketing Authorisation in the UK for this indication; not appraised by NICE but funded via the CDF)	• MabThera (rituximab) in combination with chemotherapy, followed by MabThera maintenance therapy in patients achieving a response	• Induction with MabThera monotherapy is not an appropriate comparator for patients with advanced, symptomatic FL for which the standard of care is MabThera in combination with chemotherapy. NICE guidelines recommend the use of MabThera monotherapy induction in advanced asymptomatic patients only who would not be treated with chemotherapy but may be managed by observation ('watch and wait').	No evidence has been presented for: • Rituximab monotherapy • Rituximab-based chemotherapy, without rituximab maintenance treatment • Bendamustine monotherapy

	Final scope issued by NICE	Decision problem addressed in the company submission	Companies rationale if different from the final NICE scope	ERG comment
			• Wording on MabThera use aligned with use in current clinical practice • SACT and market research data indicates little use of bendamustine as monotherapy. Bendamustine is considered only in combination with MabThera in the first-line FL induction setting	
Outcomes	The outcome measures to be considered include: • overall survival • progression-free survival • overall response rate • adverse effects of treatment • health-related quality of life	The outcome measures to be considered include: • overall survival • progression-free survival • overall response rate • adverse effects of treatment • health-related quality of life	No difference	In line with the scope of the decision problem, although ORR is not explicitly part of the economic model.
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared.	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared.	No difference	In line with the scope of the decision problem. The cost effectiveness of treatments was expressed in terms of cost per quality-adjusted life year gained. The time horizon was 40 years in the original model and 50 years in the version after clarification. An NHS and Personal Social Services perspective was adopted.

	Final scope issued by NICE	Decision problem addressed in the company submission	Companies rationale if different from the final NICE scope	ERG comment
	Costs will be considered from an NHS and Personal Social Services perspective. The availability of any patient access schemes for the intervention or comparator technologies will be taken into account. The availability and cost of biosimilar products should be taken into account.	Costs will be considered from an NHS and Personal Social Services perspective. The availability of any patient access schemes for the intervention or comparator technologies will be taken into account. The availability and cost of biosimilar products should be taken into account.		
Subgroups to be considered	None	None	No difference	In line with the scope of the decision problem.
Special considerations including issues related to equity or equality	None identified	None identified	No difference	

Source: CS, Table 1, page 10-12.[1]

CDF = Cancer Drug Fund; CHOP = Cyclophosphamide, doxorubicin, vincristine and prednisone; CVP = Cyclophosphamide, vincristine, and prednisone; FL = Follicular lymphoma; SACT = Systemic Anti-Cancer Therapy (chemotherapy dataset).

3.1 Population

The population defined in the scope is: "People with untreated advanced follicular lymphoma".

The anticipated indication for obinutuzumab is: "Gazyvaro in combination with CHOP, CVP or Bendamustine, followed by Gazyvaro maintenance therapy in patients achieving a response, is indicated for the treatment of patients with previously untreated advanced follicular lymphoma except for FL grade 3b". However, this may be modified following comments from the CHMP.

Apart from the exclusion of FL grade 3b, the population is in line with the scope. The main trial for this submission, the GALLIUM trial did not include patients with FL grade 3b.

The UK lymphoma Association describes grade 3b as: "Grade 3b lymphomas are likely to grow faster than the other grades of follicular lymphoma. In fact, grade 3b lymphomas behave more like a high-grade (fast-growing) lymphoma than a low-grade lymphoma. They are usually treated in the same way as diffuse large B-cell lymphoma (DLBCL), a type of high-grade lymphoma."[9]

3.2 Intervention

The intervention is in line with the scope. The intervention described in the scope is 'Obinutuzumab in combination with chemotherapy, with or without obinutuzumab maintenance therapy'. The intervention in the CS and the main trial is 'Obinutuzumab in combination with chemotherapy (CVP, CHOP or bendamustine), followed by obinutuzumab maintenance therapy in patients achieving a response'. The company does not provide any evidence for obinutuzumab induction therapy without maintenance therapy.

A marketing authorisation application for obinutuzumab in combination with chemotherapy, followed by obinutuzumab monotherapy as maintenance was submitted to the European Medicines Agency (EMA) in March 2017. An opinion from the EMA is anticipated in May 2017 and regulatory approval is expected in July/August 2017.

Obinutuzumab will be contraindicated to people who demonstrate hypersensitivity to obinutuzumab or to any of the following: L-histidine, L-histidine hydrochloride monohydrate, Trehalose dehydrate, Poloxamer 188, or water for injections.

Obinutuzumab is administered on a 28-day cycle basis in induction with chemotherapy and every two months in maintenance. In induction therapy obinutuzumab is administered on days 1, 8, and 15 of cycle 1, and day 1 of cycles 2–6 (1,000 mg by intravenous infusion). These infusions typically take place in a hospital with an established oncology unit, which has the staffing and infrastructure required for administration of cancer treatments.

The average length of a course of treatment is six to eight cycles induction followed by up to 12 maintenance doses for responders to induction therapy (i.e. one maintenance dose every two months for up to two years or until progression). A person with previously untreated FL is expected to receive only one course of induction therapy followed by maintenance for responders.

The dosing frequency is as follows:

Obinutuzumab induction in combination with chemotherapy (obin-chemo):

- With CHOP: 1,000 mg fixed dose obinutuzumab on days 1, 8, and 15 of Cycle 1 and on Day 1 of Cycles 2–8 (21-day cycles)
- With CVP: 1,000 mg fixed dose obinutuzumab on days 1, 8, and 15 of Cycle 1 and on Day 1 of Cycles 2-8 (21-day cycles)

- With bendamustine: 1,000 mg fixed dose obinutuzumab on days 1, 8, and 15 of Cycle 1 and on Day 1 of Cycles 2–6 (28-day cycles)

Obinutuzumab maintenance:

- 1,000 mg fixed dose obinutuzumab once every two months for up to two years or until progression, whichever occurs first.

3.3 Comparators

The company has presented evidence for one of the comparators in the scope: rituximab-based chemotherapy, with rituximab maintenance treatment (this is the comparator in the GALLIUM trial).

No evidence has been presented for: rituximab mono-therapy; rituximab-based chemotherapy, without rituximab maintenance treatment; and bendamustine mono-therapy.

As the anticipated indication for obinutuzumab includes maintenance therapy, it seems obvious that the comparator should also include maintenance. Therefore, rituximab-based chemotherapy without rituximab maintenance treatment can be ignored as a relevant comparator. However, the company should have presented evidence of obinutuzumab versus rituximab mono-therapy and bendamustine mono-therapy as specified in the scope.

3.4 Outcomes

The NICE final scope lists the following outcome measures:

- overall survival
- progression-free survival
- overall response rate
- adverse effects of treatment
- health-related quality of life

These outcomes are reported in the CS. However, OS data are still immature at the clinical cut-off date (31 January 2016) of the GALLIUM trial, with less than 20% of patients followed for survival for more than four years.

3.5 Other relevant factors

The submission includes a Patient Access Scheme (PAS). The PAS is a [redacted].

Obinutuzumab requires additional administration in induction in combination with chemotherapy in comparison to R-chemo. Furthermore, patients in England can be offered the subcutaneous formulation of MabThera for maintenance treatment after response to R-chemo induction, whereas obinutuzumab requires IV administration in maintenance.

No specific equity considerations have been raised by the company.

According to the company obinutuzumab is an innovative treatment because it is a first-in-class Type II glycoengineered anti-CD20 antibody with a mode of action based on enhanced antibody dependent cellular cytotoxicity, increased direct cell death, and a lower degree of complement dependent cytotoxicity compared with nonglycoengineered, Type I antibodies such as MabThera and ofatumumab (See: CS, page 31).

According to the company this technology does not meet the end-of-life criteria because patients with FL are expected to have life expectancy beyond 24 months (See: CS, page 126, Table 56).

4. CLINICAL EFFECTIVENESS

4.1 Critique of the methods of review(s)

The company conducted a systematic review to identify all published and unpublished RCT evidence on the use of obinutuzumab in previously untreated follicular lymphoma (FL). This section critiques the methods of the review including searching, inclusion criteria, data extraction, quality assessment and evidence synthesis.

4.1.1 Searches

Searching was conducted to identify relevant RCT evidence relating to the effectiveness of obinutuzumab in previously untreated follicular lymphoma.

Searches were reported for PubMed, Embase and the Cochrane Central Register of Controlled Trials (CENTRAL), and were undertaken in June 2015. Update searches were reported for March 2017. In order to address the limitations queried by the ERG in the clarification process, further revised searches were carried out in June 2017. The date span for each database was not reported, and strategies included a date restriction of 1998 onwards.

Supplementary searches were carried out in five conference proceedings, four trials registers or portals, and a number of relevant organisational websites. No date of search was reported for the supplementary searches, and details of search terms were provided in response to the clarification process.[8, 10] These meet the requirements detailed in the NICE guide to the methods of technology appraisal.[11]

Search strategies for the database searches were provided in the Appendix 3 of the CS[12] and in two Excel spreadsheets supplied as part of the clarification response.[8, 10, 13]

ERG comment:

- The Canadian Agency for Drugs and Technologies in Health (CADTH) evidence based checklist for the Peer Review of Electronic Search Strategies, was used to inform this critique.[14] The submission was checked against the Single Technology Appraisal (STA) specification for company/sponsor submission of evidence.[15] The ERG has presented only the major limitations of each of the revised search strategies in the main report. Further criticisms of each search strategy can be found in Appendix 1 of this report.
- For the most part, the searches were well reported and reproducible; the names of the database hosts were provided in the clarification response.[8] The database searches were well structured and used combinations of index terms appropriate to the resource searched, free text and a number of synonyms for the condition and an RCT study design filter[16] was used to further restrict the search results.
- All clinical effectiveness searches were restricted to references with line of treatment (newly diagnosed or untreated patients) in the title or abstract. The ERG considered this overly restrictive as it is possible that a relevant study might not describe line of treatment in the title or abstract. The ERG raised this point during clarification, and the company disagreed and felt RCTs in follicular lymphoma would state line of treatment in the title or abstract. However the ERG remains concerned that restriction of all searches to line of treatment was overly restrictive and problematic. There were no appropriate indexing terms for this concept in PubMed or Embase, therefore this restriction was entirely dependent on free-text terms.
- The ERG noted in the search strategies that the RCT filter used a line to remove observational studies from the final search results (line #34).[13] Therefore not all of the non-randomised evidence may have been identified.

- No specific AE searches were performed. Guidance by the Centre for Reviews and Dissemination (CRD)[17] recommends that if searches have been limited by a study design filter, additional searches should be undertaken to ensure that adverse events that are long-term, rare or unanticipated are not missed. The ERG considered that it was possible that some relevant evidence may not have been identified as a consequence of the study design limits used. Unfortunately the ERG was unable to undertake independent AE searches and review the results within the STA timeline, as this would be outside of the ERG remit.
- Significant problems resulting from incorrect use of within-phrase wildcard characters were noted in the revised CENTRAL search, presented in the clarification response.[13] Lines #2, #3, #4, #7 and #8 all utilise the wildcard command "?" within phrase terms. The Cochrane Library search help clearly states "Phrase search does NOT support the use of wildcards".[18] This problem affected each part of the disease facet, as well as the lines restricting the search to untreated or newly diagnosed disease. The ERG did not consider the CENTRAL search adequately robust to inform the clinical effectiveness systematic review, as the only unaffected lines involve subject indexing.

4.1.2 Inclusion criteria

The eligibility criteria for the systematic review of RCTs of obinutuzumab is presented in Table 4.1.

Table 4.1: Eligibility criteria used in search strategy for RCT evidence

Domain	Inclusion criteria	Exclusion criteria
Population	Patients with previously untreated iNHL	Not focussing on human data Not iNHL Not previously untreated iNHL
Interventions and comparators	All licensed and investigative interventions	Not including treatment of interest
Outcomes	All primary and secondary outcomes available, including all efficacy, all end-points, PROs, HRQoL outcomes, and safety Examples include but are not restricted to: • Efficacy endpoints reported in studies, including PFS, ORR, OS, complete remission, complete response, partial response, EFS, MRD and others • Safety endpoints reported in studies, including AEs, serious AEs, AEs leading to death, treatment discontinuations and others • HRQoL endpoints reported, including all PROs • HRQoL cancer specific: FACT-G, Mental Adjustment to Cancer Scale	Not including the outcome of interest

Domain	Inclusion criteria	Exclusion criteria
Study design	Randomised controlled trials • Non-randomised trials, or single arm trials flagged only if the population and outcomes are of interest	Not study type of interest Not publication type of interest
Source: Table 16 of the company submission (CS)[1] AE = adverse event; EFS = event-free survival; FACT-G = functional assessment of cancer therapy-general; HRQoL = health-related quality of life; iNHL = indolent non-Hodgkin lymphoma; MRD = minimal residual disease; ORR = overall response rate; OS = overall survival; PFS = progression-free survival; PRO = patient-reported outcomes		

Two additional exclusion criteria were included to filter the search results further:

1. At least one of the treatment arms includes treatment with MabThera (rituximab)
2. Studies assessing induction and maintenance treatment phases should not have a condition of successful completion of the induction treatment for patients in order to enter maintenance phase.

ERG comment:

- The population of the review is in line with the scope but the intervention is not. Regarding interventions, only studies that included a rituximab arm were eligible. The company was asked to justify the exclusion of trials including bendamustine. They stated that '*One of the treatment arms had to include rituximab as rituximab plus chemotherapy is the standard of care for the first line-line treatment of advanced FL. The only study comparing obinutuzumab and bendamustine without rituximab is the GADOLIN study, which is in patients with rituximab-relapsed / refractory FL, therefore not relevant to the indication being appraised.*'[8]
- Randomised trials were prioritised, with non-randomised trials to be flagged for use as supplementary evidence. This approach is line with NICE requirements.

4.1.3 Critique of data extraction

In response to clarification, the company stated that 'An independent reviewer undertook the quality check of the data extraction by randomly reviewing 15% of the extracted articles. Any discrepancies were resolved by discussion and a third reviewer was consulted for unresolved disagreements. The 15% QC did not identify any major mistakes, therefore no additional QC was conducted.'[8]

ERG comment:

- When conducting systematic reviews it is normally recommended that two reviewers are involved in all data extraction to reduce the potential for error and bias.

4.1.4 Quality assessment

Quality was assessed for both the included RCT and the non-RCT according to established guidance. This included assessment of randomisation, concealment of allocation of treatment, blinding of patients, care providers and assessors, trial drop-out, reporting of all measured outcomes and use of intention to treat analysis.

ERG comment:

- As above, it is normally recommended that two reviewers are involved in all quality assessment to reduce the potential for error and bias. The ERG has assessed the quality of the included trial, GALLIUM, and results are presented in section 4.2.2.4.

4.1.5 Evidence synthesis

No meta-analysis or indirect comparison could be performed as only one trial was found eligible for inclusion in the submission.

4.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

4.2.1 Overview of the evidence in the submission

The CS was based on one trial (GALLIUM) which will be discussed in detail in this section. One non-randomised trial, GAUDI, was presented as supporting evidence in the CS. The non-randomised trial is discussed briefly in section 4.2.3. No ongoing trials, other than GALLIUM were identified. The implications of the ongoing analysis of GALLIUM are discussed in section 4.2.4.

ERG comment:

- The ERG was provided with a list of excluded studies.[8] We checked the list and found no relevant RCT evidence that was omitted from the review.
- The ERG queried whether GAUDI was the only non-randomised trial of relevance to the decision problem particularly in terms of adverse events. The company clarified that this was the case.[8] The company stated that they had provided the ERG with a bibliographical list of the non-RCTs but the list could not be located. Therefore the ERG could not verify the exclusion of other non-RCTs.

4.2.2 The GALLIUM trial

4.2.2.1 Methodology of the GALLIUM trial

According to the CS '*Gallium is an ongoing Phase III, open-label, multicentre, randomised study to investigate the efficacy and safety of G-chemo followed by G maintenance monotherapy for responders (complete response [CR] or partial response [PR], compared with R-chemo followed by R-maintenance therapy for responders, in patients with previously untreated advanced indolent NHL requiring treatment*'.[1]

The GALLIUM trial included 1,401 previously untreated adult patients who had CD20-positive iNHL. GALLIUM was conducted at 177 trial centres in 18 countries. The CS stated that 293 patients (21%) of patients were from the UK. Details of inclusion and exclusion criteria were provided by the company. The methodology of the trial is summarised in Table 4.2.

Table 4.2: Methodology of the GALLIUM trial

PICOS	Details
Population	• Previously untreated CD20-positive iNHL • FL (grade 1 to 3a) or splenic/nodal/extranodal MZL • Stage III/IV or stage II bulky disease (≥ 7cm) requiring treatment • Aged ≥ 18 years • ECOG 0 to 2
Intervention	Obin-chemo as induction followed by Obin maintenance monotherapy

PICOS	Details
Comparator	R-chemo as induction followed by R maintenance monotherapy
Outcomes	Induction period only: Complete response and End-of-treatment overall response All study periods (induction + maintenance + follow up): Progression-free survival (Investigator and IRC), Overall survival, Best overall response, Disease-free survival, Event-free survival, Duration of response, Patient-reported outcomes (FACT-Lym and EQ-5D), Minimal residual disease, End-of-maintenance response, Adverse events
Study design	Randomised Controlled Trial (RCT)

Source: Tables 17 and 22 of the CS[1]
ECOG = Eastern Cooperative Oncology Group; FL = follicular lymphoma; iNHL = indolent non-Hodgkin lymphoma; IRC = independent review committee; MZL = marginal zone lymphoma; Obin-chemo = obinutuzumab with chemotherapy as induction, R-chemo = rituximab with chemotherapy as induction,

Of the 1,401 participants randomised, 1,202 (86%) had follicular lymphoma and form the basis of the submission and the primary efficacy ITT population in the trial. The remainder of this report focuses on the patients with FL.

Patients were randomly assigned by an interactive voice system (IVRS) in a 1:1 ratio to either obin-chemo followed by obin-maintenance in responders, or R-chemo followed by R-maintenance in responders. Stratification factors for randomisation were: chemotherapy regimen (CHOP, CVP or bendamustine); FLIPI (low or high for FL); and geographic region (Western Europe, Eastern Europe, South and Central America, North America, other). FL and Marginal zone lymphoma (MZL) patients were randomised separately.

Prior to the initiation of the study, each site chose one of three chemotherapy regimens (CHOP, CVP, or bendamustine) that was considered to be the standard of care for follicular lymphoma; all patients with follicular lymphoma at that site received the chosen chemotherapy regimen for the duration of the study.

In the obin-chemo arm, eight to 10 doses of obin at 1,000 mg were administered by IV infusion with the accompanying chemotherapy regimen during induction.

- **G-CHOP:** G was administered on Days 1, 8, and 15 of Cycle 1 and on Day 1 of Cycles 2–8 (21-day cycles). CHOP was administered on Day 1, with prednisone/prednisolone/methyl-prednisolone also administered on Days 2–5 of Cycles 1–6;
- **G-CVP:** G was administered on Days 1, 8, and 15 of Cycle 1 and on Day 1 of Cycles 2–8 (21-day cycles). CVP was administered on Day 1, with prednisone/prednisolone/methylprednisolone also administered on Days 2–5 of Cycles 1–8;
- **G-bendamustine:** G was administered on Days 1, 8, and 15 of Cycle 1 and on Day 1 of Cycles 2–6 (28-day cycles). Bendamustine was administered on Days 1 and 2 of Cycles 1–6, with prednisone/prednisolone/methylprednisolone administered on Day 1 of Cycle 1.

Patients randomised to receive obin-chemo who achieved a CR or PR at the end of induction therapy continued to receive obin-maintenance at 1000 mg every two months until disease progression, or for two years.

In the R-chemo arm, six to eight doses of R at 375 mg/m2 were administered by IV infusion with the accompanying chemotherapy regimen during induction.

- **R-CHOP:** R was administered on Day 1 of Cycles 1–8 (21-day cycles). CHOP was administered on Day 1, with prednisone/prednisolone/methylprednisolone also administered on Days 2–5, of Cycles 1–6;
- **R-CVP:** R was administered on Day 1 of Cycles 1–8 (21-day cycles). CVP was administered on Day 1, with prednisone/prednisolone/methylprednisolone also administered on Days 2–5, of Cycles 1–8;
- **R-bendamustine:** R was administered on Day 1 of Cycles 1–6 (28-day cycles). Bendamustine was administered on Days 1 and 2 of Cycles 1–6, with prednisone/prednisolone/methyl-prednisolone also administered on Day 1 of Cycle 1.

Patients randomised to receive R-chemo who achieved a CR or PR at the end of induction therapy continued to receive R-maintenance at 375 mg/m2 every two months until disease progression, or for two years. A modified version of the Revised Response Criteria was used to ascertain response.[19]

Dose reductions were not recommended.

Following the completion of induction therapy, patients received maintenance therapy (if they achieved a CR or PR) or underwent observation (patients with stable disease [SD]), and were followed clinically every two months for two years. For patients who had not progressed at the maintenance or observation completion visit, disease assessments continued every three months for 3threeyears then every six months for two years until disease progression. After five years of follow-up or disease progression (whichever came first), patients were then followed every six months for OS and new anti-lymphoma treatment (NALT), or for disease progression if applicable, until the end of the study, which is estimated as 10.2 years after inclusion of the first patient. Patients who terminated early without PD were followed for PD, and in the extended follow-up for PD, NALT and OS. Patients who terminated induction early because of PD went directly into the extended follow-up for NALT and OS. Patients who discontinued the protocol-defined treatment path and needed to start a NALT in the absence of disease progression (e.g., if wrong diagnosis at screening and new diagnosis required a change of treatment) were followed for disease progression and OS.

The primary outcome of GALLIUM was investigator-assessed progression-free survival (PFS). Secondary outcomes included IRC-rated PFS and overall survival (OS), response rates at induction, maintenance and follow-up. Health-related quality of life was evaluated using the disease-specific Functional Assessment of Cancer Therapy for Patients with Lymphoma (FACT-Lym) instrument and EQ-5D summary scores as follows:

- Change from baseline in all domains of the FACT-G
- Change from baseline in the total outcome index (TOI) (range, 0−116): sum of physical well-being (seven items), functional well-being (seven items), and Lym subscale (15 items) scores
- Change from baseline in the FACT-Lym subscale score (range, 0−60): 15 lymphoma-specific items
- Change from baseline in the FACT-Lym total score (range, 0−168): sum of physical well-being (seven items), social/family well-being (seven items), emotional well-being (6 items), functional well-being (seven items), and Lym subscale (15 items) scores
- EQ-5D summary scores at baseline, during treatment, after treatment, at the last assessment prior to progression, and at the first assessment after progression

After the initiation of the study medication, adverse events (AEs) and serious adverse events (SAEs) were recorded as follows (until patient began NALT):

- All AEs (related and unrelated) were recorded up to 28 days are the last dose of study drug
- Grade ≥3 AEs (related and unrelated) were recorded up to six months after the last dose of study drug
- Grade 3 or 4 infections (related and unrelated) were recorded up to 24 months after the last dose of study drug
- Unrelated SAEs were recorded up to 12 months after the last dose of study drug
- Study drug-related SAEs were recorded indefinitely (even if the study had been closed).

Pre-specified subgroup analyses were conducted for investigator-assessed PFS, IRC-assessed PFS, CR rate and ORR (all without PET) for the following:

- Stratification factors (chemotherapy regimen, FLIPI or IPI risk group, geographic region)
- Age at randomisation
- Baseline characteristics and disease demographics.

Results for subgroups are presented in Section 4.2.2.7.

ERG comment:

- The ERG notes that GALLIUM is a large, multicentre trial in a relevant population, investigating important, patient-relevant outcomes.
- The trial includes a reasonable proportion of UK patients (21%) and nearly half from Western Europe so from this perspective is relevant to the UK.
- The ERG draws to the attention of the committee that patients with grade 3b lymphoma were excluded from GALLIUM which is in line with the anticipated indication for obinutuzumab.
- GALLIUM investigates both induction and maintenance treatment with obinutuzumab. All participants who achieved a response entered maintenance. The NICE scope specified an assessment of obinutuzumab with / without maintenance. In this report we provide response rates from GALLIUM at the end of induction and fuller outcomes for induction and maintenance which reflects the intended indication of obinutuzumab.
- Three types of chemotherapy are offered to patients in the trial (CHOP, CVP and bendamustine). However the trial was not designed to investigate differences in chemotherapy regimens. The committee will need to decide if the breakdown of regimens reflects UK clinical practice. This issue is discussed further in Section 4.2.2.3.
- Independent committee outcomes are used in GALLIUM in addition to investigator outcomes. As the trial is open-label these will be more reliable.

4.2.2.2 Statistical analysis of the GALLIUM trial

Sample size calculation

The primary analysis compared PFS in the R-chemo and obin-chemo arms using of a two-sided stratified log rank test at an overall 5% significance level.

In the FL population, it was estimated that 370 PFS events were required overall to demonstrate efficacy based on the following assumptions:

- Two-sided log rank test at the 0.05 level of significance;
- 80% power to detect a hazard ratio (HR) for obin-chemo versus R-chemo of 0.74, corresponding to an improvement in three-year PFS from 70.7% to 77.4% or in median PFS from six years to 8.1 years (35%). Estimates of median PFS were not likely to be reached in either study arm at either interim or final PFS analysis;

- Exponential distribution of PFS;
- An annual dropout rate of 2.5%.

Analysis methods

Three interim analyses were planned, two for futility (one for CR and one for PRS) and one for efficacy (for PFS). The first interim analysis was performed after the first 170 patients with FL and the Independent Data Monitoring Committee (IDMC) recommended that the study continue. The second interim analysis was conducted when approximately 111 PFS events had occurred and the IDMC also recommended continuation of the study. The third interim analysis for efficacy was performed after approximately PFS events using a data cut-off of 31 January 2016. The IDMC reviewed the data on 20 May 2016 and recommended that the study be fully analysed at this time, as the primary endpoint had been met. This is the primary analysis in the CS. The ERG also received full data from the company for analyses using the 10 September 2016 clinical cut-off date.

The trial has four main analysis populations:

- **ITT – FL population:** The primary efficacy analysis population is the ITT FL population, defined as all randomised patients with follicular histology. Efficacy analyses were conducted according to the ITT principle, where patients were grouped according to their randomised treatment arm regardless of what treatments were actually received.
- **ITT – overall population:** The primary and key secondary efficacy parameters were also determined in the overall ITT population, defined as all randomised patients.
- **Safety Population:** The safety analysis population included all patients who received any amount of study drug (Obin, R, or chemotherapy [CHOP, CVP, or bendamustine]), and patients were analysed according to the treatment received (i.e., a patient who received obin at least once for any reason was analysed under the obin-chemo treatment arm; if only chemotherapy and/or R was received, the patient was analysed under the R-chemo treatment arm).
- **PET evaluable population:** The "PET evaluable" subset contains all patients for whom the answer to the question "Were there any PET-avid lesions representing lymphoma?" on PET scan eCRF at baseline was "Yes".

Results from all four analysis populations were available in the submission.

PFS was the primary efficacy endpoint of GALLIUM, defined as the time from the day of randomisation until the first documented day of disease progression, symptomatic deterioration, disease transformation, or death from any cause, whichever occurred first. Patients who did not experience documented disease progression or death were censored at the last valid (SD, PR, CR) tumour assessment prior to the clinical cut-off date.

PFS was compared using a two-sided log-rank test stratified by chemotherapy regimen (CHOP, CVP, or bendamustine), FL international prognostic index (FLIPI) risk group (low, intermediate, or high) in patients with FL or international prognostic index (IPI) risk group (low or low-intermediate vs. high-intermediate or high) in patients with non-follicular lymphoma. Estimates of the treatment effect, reported as hazard ratios (HR) with 95% confidence intervals (CI) were obtained from a stratified Cox model. An unstratified log-rank test was performed as a sensitivity analysis. Estimates of two and three-year survival with 95% CI for each treatment arm were obtained using the Kaplan-Meier method.

To adjust for multiple statistical testing and control the overall Type 1 error rates at a two-sided 0.05 significance level a fixed sequence testing procedure was used. Endpoints were tested in the following order:

- PFS in the overall population
- CR rate at the end of induction therapy in the FL population based on tumour assessment without PET
- CR rate at the end of induction therapy in the overall population based on tumour assessment without PET
- Overall survival in the FL population
- Overall survival in the overall population
- ORR at the end of induction therapy in the FL population based on tumour assessment without PET
- ORR at the end of induction therapy in the overall population based on tumour assessment without PET

All analyses were based on the investigator's assessment. PFS, CR and ORR were based on IRC assessments for US registration purposes.

ERG comment:

- The statistical aspects of the study design including the sample size calculation and interim analyses were appropriate. The statistical analysis methods also seem to be appropriate.

4.2.2.3 Participants in the GALLIUM trial

A total of 1,202 FL patients were randomised in the study (601 patients to the R-chemo arm and 601 patients to the obin-chemo arm).

The overall median observation time (randomisation to last available assessment) at the cut-off date (January 2016) was 34.4 months (range: 0.1–54.5 months) in the R-chemo arm and 34.8 months (range: 0.0–53.8 months) in the obin-chemo arm. The proportion of patients who had been observed for at least two years at the clinical cut-off was 87.7% in the R-chemo arm and 91.3% in the obin-chemo arm. At the clinical cut-off date, 44.1% of patients in the R-chemo arm and 45.1% of patients in the obin-chemo arm had been followed for at least three years.

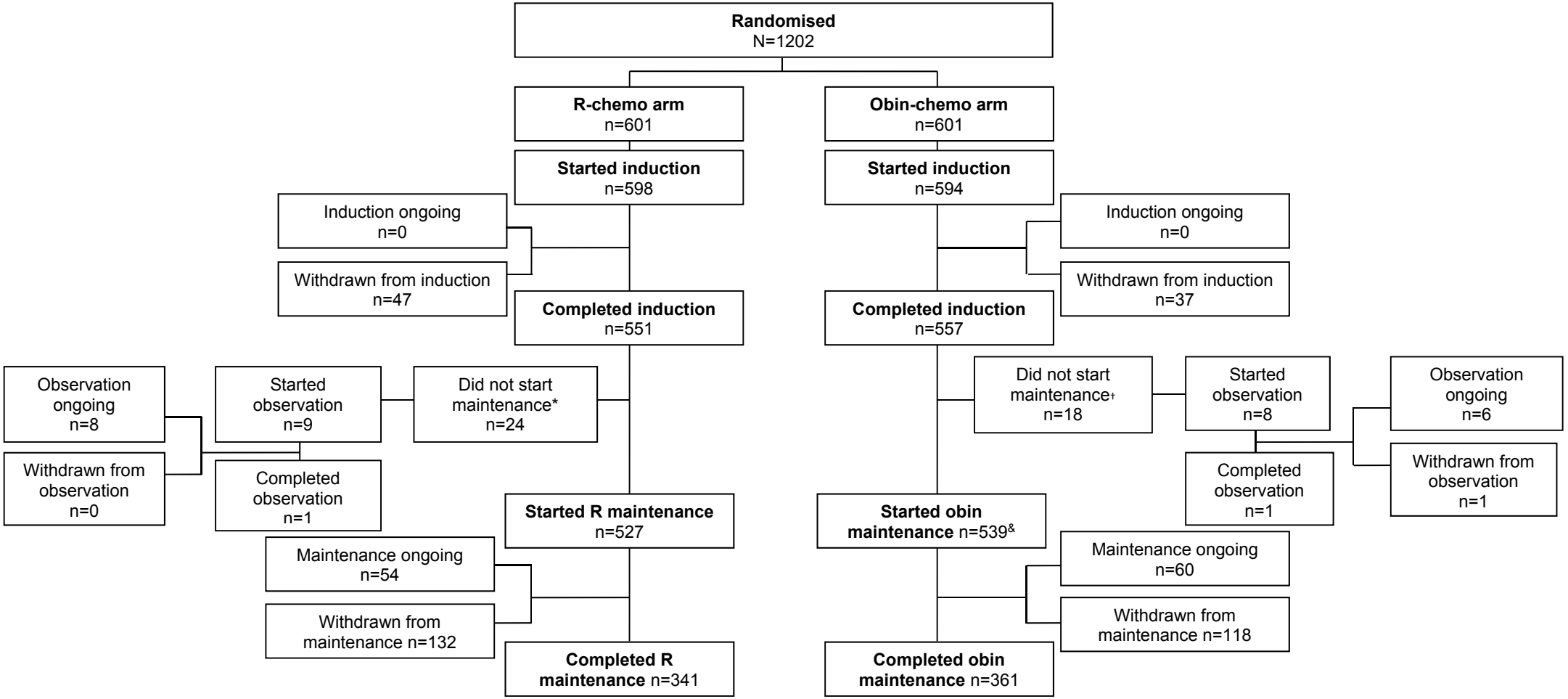
The median duration of post-treatment follow-up at the cut-off date was 9.2 months (range: 0.0–42.3 months) in the R-chemo arm and 9.4 months (range: 0.0–46.9 months) in the obin-chemo arm.

During the induction phase, 7.8% patients in the R-chemo arm and 6.2% patients in the obin-chemo arm of the FL population were withdrawn from treatment. Most withdrawals were due to AEs and comparable between treatment arms.

During the maintenance phase, 22.0% patients in the R-chemo arm and 19.6% patients in the obin-chemo arm of the FL population were withdrawn from treatment. The main reason for withdrawals was progressive disease with a higher proportion of patients in the R-chemo arm (10.6% compared with 6.2% in the obin-chemo arm).

A participant flow diagram for the GALLIUM trial as of the data cut-off date for the interim analysis (31 January 2016) is provided in Figure 4.1.

Figure 4.1: Participant flow in GALLIUM (cut-off 31 January 2016)



Source: CS, Figure 9, page 75[1] with corrections from response to letter of clarification.[8]

*24 patients did not start R-maintenance treatment due to: progressive disease between induction and maintenance (n=10); started observation (i.e., stable disease) (n=9); withdrawal by subject (n=3); physician decision (n=1); and other (n=1). †19 patients did not start G-maintenance treatment due to: progressive disease between induction and maintenance (n =10; started observation (i.e., stable disease) (n=8); and withdrawal by subject (n=1) & 1 additional patient entered maintenance without completing induction

In the FL population, the median age of patients was 59 years (range: 23–88 years); overall, more female than male patients were randomised (53.2% vs. 46.8%). The trial population was predominantly Caucasian (80%) with just four black or African American participants. As previously stated across the whole trial 21% were from the UK but in the follicular lymphoma population specifically approximately 48% were from Western Europe.

The overall median time from first diagnosis to randomisation was 1.5 months (range: 0.0–168.1 months). Mean time from diagnosis was 6.25 months in the obinutuzumab group and 7.28 in the rituximab group. Of the three different chemotherapy regimens permitted the most frequently used was bendamustine (57%), then CHOP (33%) and finally CVP (10%).

The majority of patients had an ECOG performance status of 0-1 (96.8%). The greatest proportion of patients comprised intermediate and high-risk FLIPI (37.2% and 41.8% respectively) and FLIPI-2 groups (50.3% and 40.6%, respectively), and Ann Arbor stage III-IV (>91%). Nearly half (43.8%) of patients had a nodal or extra-nodal mass over 7 cm in diameter. There was extra-nodal involvement in 65.6% of patients.

Demographic and clinical characteristics of patients enrolled in GALLIUM are summarised in Table 4.3.

Table 4.3: Patient demographics and baseline characteristics (ITT population)

Domain	**O-chemo (n = 601)**	**R-chemo (n = 601)**
Mean age, years (SD)	58.2 (11.5)	57.7 (12.2)
Male, n (%)	283 (47.1)	280 (46.6)
Mean height, cm (SD)	168.3 (10.0)	168.4 (10.1)
Mean weight, kg (SD)	76.3 (17.9)	75.2 (17.0)
Mean body surface area, m2 (SD)	1.86 (0.2)	1.84 (0.2)
Mean BMI, kg/m2 (SD)	26.8 (5.3)	26.4 (5.9)
Race, n (%) Caucasian Black or African American Asian American Indian or Alaska Native Native Hawaiian or other Pacific islander Multiple Other	 487 (81.0) 3 (0.5) 100 (16.6) 0 1 (0.2) 0 10 (1.7)	 481 (80.0) 1 (0.2) 98 (16.3) 1 (0.2) 0 3 (0.5) 17 (2.8)
Geographic region, n (%) Eastern Europe Western Europe North America Asia Other	 78 (13.0) 294 (48.9) 75 (12.5) 92 (15.3) 62 (10.3)	 79 (13.1) 286 (47.6) 77 (12.8) 93 (15.5) 66 (11.0)
ECOG PS, n (%) 0–1 2	n=600 585 (97.5) 15 (2.5)	n=599 576 (96.2) 23 (3.8)

Domain	O-chemo (n = 601)	R-chemo (n = 601)
Ann Arbor Stage, n (%) I II III IV	n=598 10 (1.7) 41 (6.9) 208 (34.8) 339 (56.7)	n=597 8 (1.3) 44 (7.4) 209 (35.0) 336 (56.3)
FLIPI no. of adverse factors categories 1, n (%) Low (0,1) Intermediate (2) High (≥3)	n=601 128 (21.3) 224 (37.3) 249 (41.4)	n=601 125 (20.8) 223 (37.1) 253 (42.1)
FLIPI no. of adverse factors categories 2, n (%) Low (0,1) Intermediate (2) High (≥3)	n=579 51 (8.8) 296 (51.1) 232 (40.1)	n=586 55 (9.4) 290 (49.5) 241 (41.1)
Bone marrow involvement at BL, n/patients with data (%)	318/592 (53.7)	295/598 (49.3)
Extranodal involvement, n/patients with data (%)	392/601 (65.2)	396/601 (65.9)
Bulky disease at BL (6 cm threshold), n/patients with data (%)	255/600 (42.5)	271/600 (45.2)
Mean time from diagnosis to randomisation, months (range)	6.25 (0.1–121.6)	7.28 (0.0–168.1)
Chemotherapy regimen, n (%) Bendamustine CHOP CVP	 345 (57.4) 195 (32.4) 61 (10.1)	 341 (56.7) 203 (33.8) 57 (9.5)
Source: Table 25 of the CS[1] ECOG = Eastern Cooperative Oncology Group performance score; FLIPI = follicular Lymphoma International Prognostic Index; SD = standard deviation		

ERG comment:

- Baseline characteristics appear balanced between groups in GALLIUM. Both male and female participants are included in the trial in similar numbers. In relation to the UK, almost 50% of participants are from Western Europe. There are very few black participants in the trial (4, 0.3%).
- Baseline characteristics in the GALLIUM trial may not be entirely representative for the advanced FL population in the UK. For instance, according to the company "the median age at diagnosis of FL in the UK is approximately 65 years old" (CS, page 32), while the mean age in GALLIUM is 58 years (median: 59 years). In addition, the company reported the following proportions of patients treated with immunochemotherapy regimens as first-line treatment for FL in UK clinical practice: 36% R-CVP, 29% R-bendamustine, 13% R-CHOP and 22% other (CS, Table 14, page 43). In the GALLIUM trial, 57% received bendamustine, 33% CHOP and 10% CVP (see Table 4.3).
- Despite these differences, the company states that feedback from clinical experts confirmed that the baseline characteristics of the FL patients enrolled into GALLIUM were reflective of the population seen in UK clinical practice (CS, page 124); although the company noted that the time from diagnosis to treatment is shorter compared with clinical practice. In the

clarification letter[8] we asked the company to indicate the arguments used to state that the population in GALLIUM is reflective of the UK population regarding which chemotherapy regimen (CHOP, CVP or bendamustine) was used. The company responded that "there are local variations in clinical practice with respect to chemotherapy use and therefore, the appropriate representative average use of the three chemotherapy regimens has some uncertainty." The ERG concludes that according to the company (CS, Tables 13 and 14, page 43) from the three chemotherapy regimens, R-CVP is most often used in UK clinical practice while only 10% of patients in the GALLIUM trial received CVP. As the allocation of chemotherapy was not randomised at the patient level, there may be confounding differences in baseline patient characteristics between the chemotherapy subgroups; therefore, the baseline characteristics in the GALLIUM trial may not be representative for the advanced FL population in the UK.

- We also asked the company to provide baseline characteristics of participants by type of chemotherapy received in addition to R or obin. The baseline characteristics between chemotherapy subgroups are summarised in Table 4.4. Overall, high risk patients were more likely to receive CHOP, whereas bendamustine and CVP use was more frequent among older patients and patients with more comorbidity. According to the company, this reflects the use of chemotherapy regimens in clinical practice.

Table 4.4: Baseline characteristics by chemotherapy received (ITT population)

n (%)	Benda n=686	CHOP n=399	CVP n=117
Median age, years (range)	59 (23–88)	58 (31–85)	59 (32–85)
Age ≥80 years	**23 (3.4)**	3 (0.8)	**4 (3.4)**
Male	332 (48.4)	177 (44.4)	54 (46.2)
Charlson Comorbidity Index score ≥1	**163 (23.8)**	69 (17.3)	22 (18.8)
ECOG PS 2	24 (3.5)	8 (2.0)	6 (5.1)
FLIPI high risk (≥3)	274 (39.9)	**187 (46.9)**	41 (35.0)
Bulky disease (≥7cm)	274 (39.9)	**206 (51.6)**	46 (39.3)

- Overall, the committee will need to consider how well patients and the chemotherapy treatments in GALLIUM reflect UK clinical practice.

4.2.2.4 Quality assessment of the GALLIUM trial

Quality assessment of the GALLIUM trial is described in Table 4.5.

Table 4.5: Quality assessment of the GALLIUM trial

Question	Company assessment and explanation	ERG assessment and explanation
Was randomisation carried out appropriately?	Yes	Yes
Was the concealment of treatment allocation adequate?	N/A (open label study)	Yes randomisation was performed using an interactive voice response system (IVRS).
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Yes	Yes, there do not appear to be important differences at baseline.

Question	Company assessment and explanation	ERG assessment and explanation
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)	No (open label study)	Partial – Care providers and participants were not blinded as this is an open label study. Therefore, investigator assessment was unblinded. IRC assessors were blinded, making this the more reliable assessment.
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	No	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No. The CSR was provided for both January 2016 and September 2016 cut-offs. Overall survival results are not mature.
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for the data?	Yes	Yes
Source: Table 27 of the CS[1]		

ERG comment:

- The GALLIUM trial appears to be well conducted. However it is an open label trial, which means patients and physicians are unblinded. Therefore, results based on the independent review committee (IRC) are more reliable.
- In addition, overall survival results were not mature at the time of the interim analysis (31 January 2016), with less than 20% of patients having been followed for survival for more than four years, and a total of 81 randomised patients that had died: 46/601 patients (7.7%) in the R-chemo+R arm and 35/601 patients (5.8%) in the obin-chemo+obin arm. Even PFS results were relatively immature with 24.0% and 16.8% (in the primary analysis) of patients having progressed or died in the R-chemo+R and obin-chemo+obin arm, respectively. For patients progressing post two years (late PD, after two years from treatment initiation), post progression survival (PPS) for GALLIUM was too immature, i.e. there were too few post-progression deaths.
- The GALLIUM study is ongoing. At the time of the clinical cut-off date (31 January 2016), 114 patients with FL were still undergoing maintenance treatment (54 in the R-chemo arm and 60 in the obin-chemo arm).

4.2.2.5 Efficacy results of the GALLIUM trial

The data reported in the clinical effectiveness section of the CS are those from the subgroup of patients with FL within the ITT population (i.e. 85.8% [1202/1401] of the ITT population). In this ERG report we will do the same.

The data reported in the CS were taken from the primary analysis (clinical cut-off 31 January 2016), although data (where available) from the updated analysis were also presented (clinical cut-off dated 16 September in the CS and corrected to 10 September 2016 in the response to clarification). In the clarification letter we asked whether any further data were available and received full data for the 10 September 2016 clinical cut-off date. Therefore, we will report both results together where possible.

Tables 4.6 summarises the key efficacy data for this study. The economic analyses are based on the most recent data (September 2016 cut-off).

Table 4.6: Summary of efficacy data from GALLIUM (FL ITT population)

	Primary analysis (January 2016 cut-off date)		Updated analysis (September 2016 cut-off date)	
	Obin-chemo n=601	**R-Chemo n=601**	**Obin-chemo n=601**	**R-Chemo n=601**
Progression-free survival (INV-assessed)				
Patients w/ event, n (%)	101 (16.8)	144 (24.0)	120 (20.0)	161 (26.8)
Median PFS, (95% CI), m	NE (NE to NE)	NE (47.1 to NE)	NE	NE
HR (stratified), 95% CI	0.66 (0.51 to 0.85), p=0.0012		0.68 (0.54 to 0.87), p=0.0016	
Progression-free survival (IRC-assessed)				
Patients w/ event, n (%)	93 (15.5)	125 (20.8)	108 (18.0)	141 (23.5)
Median PFS, (95% CI), m	NE (48.7 to NE)	51.2(47.1 to NE)	NE	NE
HR (stratified), 95% CI	0.71 (0.54 to 0.93), p=0.0138		0.72 (0.56 to 0.93), p=0.0118	
Overall survival				
Patients w/ event, n (%)	35 (5.8%)	46 (7.7%)	43 (7.2%)	52 (8.7%)
Median OS, months	NE	NE	NE	NE
HR (stratified), 95% CI	0.75 (0.49 to 1.17), p=0.21		0.82 (0.54 to 1.22), p=0.32	
Event-Free Survival				
Patients w/ event, n (%)	112 (18.6%)	159 (26.5%)	130 (21.6%)	179 (29.8%)
HR (stratified), 95% CI	0.65 (0.51 to 0.83), p=0.0006		0.66 (0.53 to 0.83), p=0.0004	
Time to New Anti-Lymphoma Treatment				
Patients w/ event, n (%)	80 (13.3%)	111 (18.5%)	86 (14.3%)	120 (20.0%)
HR (stratified), 95% CI	0.68 (0.51 to 0.91), p=0.009		0.68 (0.52 to 0.90), p=0.007	
Disease-Free Survival				
Patients incl. in analysis, n	298	281	307	293
Patients w/ event, n (%)	27 (9.1%)	33 (11.7%)	34 (11.1%)	40 (13.7%)
HR (stratified), 95% CI	0.81 (0.48 to 1.35)		0.82 (0.52 to 1.31)	
Duration of response				
Patients incl. in analysis, n	571	567	569	566
Patients w/ event, n (%)	88 (15.4%)	124 (21.9%)	105 (18.5%)	141 (24.9%)
HR (stratified), 95% CI	0.66 (0.50 to 0.87)		0.69 (0.53 to 0.88)	
Overall response (CR, PR) at end-of-induction				
Without PET, n (%)	532 (88.5%)	522 (86.9%)	530 (88.2%)	519 (86.4%)
Δ 95% CI	1.7% (-2.1 to 5.5), p=0.33		1.8% (-2.02 to 5.68), p=0.30	
With PET n (%)	N=297 255 (85.9%)	N=298 243 (81.5%)	N=297 254 (85.5%)	N=298 242 (81.2%)
Δ 95% CI	4.3% (-1.8 to 10.4), p=0.19		4.3% (-1.8 to 10.5), p=0.17	
Complete response at end-of-induction				
Without PET, n (%)	117 (19.5%)	143 (23.8%)	112 (18.6%)	145 (24.1%)
Δ 95% CI	-4.3% (-9.1 to 0.4), p=0.07		-5.5% (-10.2 to -0.78), p=0.02	
With PET n (%)	N=297 185 (62.3%)	N=298 169 (56.7%)	N=297 184 (62.0%)	N=298 169 (56.7%)
Δ 95% CI	5.6% (-2.5 to 13.6), p=0.28		5.2% (-2.8 to 13.3), p=0.32	

Source: Table 2, Response to clarification letter[8]

CI = confidence interval; CR = complete response; HR = hazard ratio; IRC = Independent Review Committee; m = months; NE = not estimable; OS = overall survival; PET = positron-emission tomography; PFS = progression-free survival; PR = Partial response

Note: Median follow up primary analysis: 34.5 months; median follow up updated analysis: 41.1 months

As can be seen from Table 4.6, results demonstrated superior PFS with obin-chemo compared with R-chemo for both cut-off dates and as per investigator and IRC assessment. OS results favoured obin-chemo, but were not statistically significant. However, results from the independent review committee (IRC) for PFS were less favourable for obin-chemo than those based on investigator assessment. Similarly, more recent results (September 2016 cut-off) were in most cases less favourable when compared with those from the January 2016 cut-off.

Other outcomes, such as: event-free survival, time to new anti-lymphoma treatment, and duration of response significantly favoured obin-chemo over R-chemo. However outcomes such as: disease-free survival, overall response and complete response (with PET) at end of induction showed no significant differences between treatments.

ERG comment:

In the ERG base case of the economic model we will use the most recent data based on IRC-assessment (i.e.: PFS=0.72 (0.56 to 0.93) (IRC, Sep 2016) instead of HR = 0.68 (0.54 to 0.87) used in the company's base-case. This is both the most up-to-date and reliable figure as it is based on IRC-assessment.

As mentioned before, OS data were immature, even at the updated clinical cut-off date (10 September 2016), with 95 patients (7.9% of the FL population) who had died, and less than 20% of patients who had been followed for survival for more than four years. The NICE committee will need to decide the nature of the relationship between improved PFS and OS.

4.2.2.6 HRQoL results of the GALLIUM trial

Patients assessed their health-related quality of life (HRQoL) using two self-administered questionnaires: the Functional Assessment of Cancer Therapy for Lymphoma (FACT-Lym) and European Quality of Life (EuroQol) EQ-5D-3L. Higher scores represent better functioning, HRQoL, and health status on both questionnaires. Change from baseline to the end of study on FACT-Lym was investigated. EQ-5D scores were assessed at baseline, during treatment, after treatment, at the last assessment prior to progression and at the first assessment after progression.

The proportions of patients randomised to each treatment arm who completed all scales on the FACT-Lym and EQ-5D questionnaires were generally balanced between treatment arms. Mean baseline scores for each of the individual FACT-Lym questionnaire subscales, and of composite FACT-G, TOI and Total scores, as well as of EQ-5D-3L utility scales were similar between R-chemo+R and obin-chemo+obin treatment arms. Both arms exhibited some impairment in the functioning and lymphoma symptom subscales as noted by mean scores of between 5 and 15 points lower than the maximum possible depending on the subscale.

There were no notable differences between the treatment arms in any of the FACT-Lym questionnaire subscales or EQ-5D-3L scales over time during the induction and maintenance treatment periods, and follow-up, as evidenced by modest (<5%) between arm differences in the mean changes from baseline scores in FACT-Lym subscales, TOI and Total score, and EQ-5D-3L utility scales.

Similar proportions of patients in the obin-chemo+obin and R-chemo+R arms had improvement in their FACT-Lym questionnaire scores during treatment and throughout maintenance and follow-up as defined by a ≥3 point increase from baseline in the Lymphoma subscale, a ≥6 point increase from baseline in the FACT Lym TOI and a ≥7 point increase from baseline in the FACT Lym Total score.

The company provided a summary of meaningful improvement in FACT-Lym. This is shown in Table 4.7. There were no statistical comparisons between the treatment groups. EQ-5D results at 36 months follow up for patients who entered the follow up phase are reported in Table 4.8.

Table 4.7: Summary of meaningful improvement in FACT-Lym in GALLIUM

FACT-Lym Subscale (definition of meaningful improvement), n (%)[a]	Obin-chemo + Obin (n = 601)	R-chemo + R (n = 601)
Lymphoma subscale (≥ 3 point increase)		
Cycle 3, Day 1 (Induction treatment)	229 (45.1)	217 (40.8)
End of Induction visit	233 (47.0)	238 (47.6)
Maintenance visit Month 2	233 (57.4)	212 (56.5)
Maintenance visit Month 12	227 (53.7)	216 (56.1)
Maintenance Completion visit	218 (56.2)	205 (55.0)
FACT TOI (≥ 6 point increase)		
Cycle 3, Day 1 (Induction treatment)	162 (31.7)	163 (30.5)
End of Induction visit	189 (38.0)	203 (40.0)
Maintenance visit Month 2	192 (47.1)	182 (48.3)
Maintenance visit Month 12	202 (47.6)	190 (49.1)
Maintenance Completion visit	191 (49.1)	174 (46.4)
FACT Total (≥ 7 point increase)		
Cycle 3, Day 1 (Induction treatment)	173 (33.9)	179 (33.5)
End of Induction visit	197 (39.6)	206 (40.6)
Maintenance visit Month 2	191 (46.8)	180 (47.7)
Maintenance visit Month 12	197 (46.5)	188 (48.5)
Maintenance Completion visit	191 (49.1)	171 (45.5)

Source: Table 40 of the CS[1]
a: Percentages are calculated on the number of patients who completed the questionnaire at each visit

Table 4.8: EQ-5D Follow-up phase (patients who entered follow-up phase) in GALLIUM

Follow up month 36	Obin-chemo + Obin (n = 453)		R-chemo + R (n = 446)	
	Value	Change from baseline	Value	Change from baseline
N	161	151	156	149
Mean (SD)	0.85 (0.22)	0.05 (0.25)	0.85 (0.20)	0.04 (0.23)
Median	1	0.04	0.85	0
Min to Max	-0.2 to 1.0	-0.8 to 1.0	-0.1 to 1.0	-0.9 to 0.7

Source: P2978 of the CSR[20]

ERG comment:

- The ERG notes that HRQoL as measured by FACT-LYM and EQ-5D were similar for each treatment group. Similar proportions had improvements in the disease-specific FACT-LYM questionnaire scores during treatment, throughout maintenance and follow-up. However, as far as we can see, there were no statistical comparisons between the treatment groups.

4.2.2.7 Subgroup analyses of the GALLIUM trial

The company assessed the potential impact of baseline demographics, prognostic factors, and stratification factors on the treatment effect. Hazard ratios for investigator-assessed PFS in the FL ITT population with 95% confidence intervals (obin-chemo+obin vs. R-chemo+R) for pre-specified patient subgroups are reported in several forest plots (CS, Figures 16 to 18, pages 98-100). Subgroups included FLIPI (low, intermediate or high), chemotherapy regimen (CHOP, CVP or bendamustine), geographic region, age, gender, race, presence of bulky disease at baseline, ≥ 1 B symptoms at baseline, Ann Arbor stage, ECOG (0 to 1 or 2) and ADL at baseline.

The company concludes that overall, the results of the PFS subgroup analyses are consistent with the primary analysis of PFS in the FL population which demonstrated improved results for obinutuzumab. They also state that the results of the IRC-assessed PFS subgroup analyses are consistent with the overall analysis of IRC-assessed PFS and with the investigator-assessed PFS subgroup analysis.

ERG comment:

- The ERG noted that there were differences in PFS according to gender. The HR for males was 0.82 (0.59 to 1.15) and for females 0.49 (0.33 to 0.74), p = 0.056.
- The majority of investigators in the GALLIUM trial chose bendamustine (57%) and <10% of investigators chose CVP as the backbone chemotherapy regimen for patients at their site. In section 4.2.2.3 of this report we mentioned that this may not be reflective of UK practice. In the UK CVP is most often chosen for patients treated with first-line immunochemotherapy regimens for FL and bendamustine is less often chosen (based on data from the company) as it is in the GALLIUM trial. Results from GALLIUM show that results for obin-chemo+obin are most favourable for patients treated with bendamustine: The observed hazard ratios by chemotherapy subgroup were as follows; CHOP (n=398): HR 0.77 (95% CI: 0.50 to 1.20), CVP (n=118): HR 0.63 (95% CI: 0.32 to 1.21), and bendamustine (n=686): HR 0.61 (95% CI: 0.43 to 0.86).
- The company states that "*subgroup analyses for the different chemotherapy regimens should be interpreted with caution because the trial was not designed to compare the efficacy of chemotherapy. The induction regimen was chosen on a per centre basis for patients with FL. Accordingly, there could be differences in patient populations treated with the different regimens.*" The company provided details of the baseline characteristics according to chemotherapy regimen and we confirmed that this was indeed the case.
- The committee needs to consider the uncertainty of the differing results for obinutuzumab according to chemotherapy regimen. The results could be a reflection of differing effects of chemotherapy regimen or that of patient selection. How closely the committee believes that GALLIUM reflects UK practice in terms of patient characteristics and chemotherapy breakdown impacts on the determination of effectiveness.

4.2.2.8 Safety results of the GALLIUM trial

The company presented data from the FL safety analysis population (i.e. patients with FL who received any amount of study drug) from the primary analysis of the GALLIUM study (clinical cut-off 31 January 2016). Overall safety results are presented in Table 4.9 below.

Table 4.9: Overall safety results in GALLIUM (31 January 2016 cut-off)

N, %	Obin-chemo + Obin (n = 595)	R-chemo + R (n = 597)
No. of patients with at least one AE (any Grade)	592 (99.5)	587 (98.3)
Total no. of events	10,311	9,343
Total no. of deaths	35 (5.9)	46 (7.7)
No. of patients with at least one AE		
AE with fatal outcome	24 (4.0)	20 (3.4)
Grade 3–5 AE	444 (74.6)	405 (67.8)
SAE	274 (46.1)	238 (39.9)
SAE leading to treatment withdrawal	44 (7.4)	36 (6.0)
SAE leading to dose reduction	12 (2.0)	10 (1.7)
SAE leading to dose interruption	83 (13.9)	45 (7.5)
Related SAE	152 (25.5)	122 (20.4)
AE leading to treatment withdrawal	97 (16.3)	85 (14.2)
AE leading to dose reduction	107 (18.0)	95 (15.9)
AE leading to dose interruption	395 (66.4)	338 (56.6)
Related AE	564 (94.8)	547 (91.6)
Related AE leading to treatment withdrawal	75 (12.6)	65 (10.9)
Related AE leading to dose reduction	103 (17.3)	89 (14.9)
Related AE leading to dose interruption	349 (58.7)	296 (49.6)

Source: Table 46 of CS[1]

Footnote: 'treatment' refers to any treatment.

AE = adverse event; obin-chemo+obin = obinutuzumab + chemotherapy followed by obinutuzumab maintenance; R-chemo+R = rituximab + chemotherapy followed by rituximab maintenance; SAE = serious adverse event

A total of 1,192 patients with FL received any amount of study drug during the induction phase (597 patients in the R-chemo arm, and 595 patients in the obin-chemo arm), and are included in the FL safety population.

During induction, most patients received all planned doses of obinutuzumab or rituximab. The median duration of treatment with rituximab and obinutuzumab during induction was the same in the two arms (25.1 weeks).

As summarised below, 526 patients in the R-chemo+R arm received R-maintenance treatment, and 540 patients in the obin-chemo+obin arm received obin-maintenance treatment. At the time of the clinical cut-off date, 114 patients with FL were still ongoing with maintenance treatment (54 in the R-chemo arm and 60 in the obin-chemo arm). The median duration of treatment with rituximab and obinutuzumab during maintenance was the same in the two arms (92 weeks).

The incidence of AEs over the entire study period (i.e., induction, maintenance and follow-up) was similar in the two treatment arms; 99.5% had at least one AE in the obin-chemo+obin arm compared with 98.0% in the R-chemo+R arm. Although, nearly all adverse events were more often reported in

the obin-chemo+obin arm than in the R-chemo+R arm. The most frequently affected System Organ Classes were as follows (percentages expressed as obin-chemo+obin vs. R-chemo+R):

- Gastrointestinal disorders (79.3% vs. 75.2%)
- Infections and infestations (77.3% vs. 70.0%)
- General disorders and administration site conditions (74.5% vs. 68.8%)
- Injury, poisoning and procedural complications (63.9% vs. 55.1%)
- Blood and lymphatic system disorders (58.3% vs. 52.8%).

The five most frequently reported AEs were (obin-chemo+obin vs. R-chemo+R):

- Infusion-related reactions (IRRs) (59.0% vs. 48.9%),
- Nausea (46.9% vs. 46.6%)
- Neutropenia (48.6% vs. 43.6%)
- Fatigue (36.0% vs. 36.5%)
- Constipation (35.3% vs. 31.5%).

Treatment-related AEs were observed in 94.8% of patients in the obin-chemo+obin arm and 91.6% of patients in the R-chemo+R arm. Related AEs were most frequently reported in the following System Organ Classes (Obin-chemo+obin vs. R-chemo+R):

- Gastrointestinal disorders (65.2% vs. 62.0%)
- General disorders and administration site conditions (60.8% vs. 50.8%)
- Injury, poisoning and procedural complications (59.2% vs. 49.1%)
- Blood and lymphatic system disorders (54.3% vs. 48.2%).

Overall, there was a higher incidence of SAEs in the obin-chemo+obin arm than in the R-chemo+R arm. A total of 274/595 patients (46.1%) in the obin-chemo arm experienced 590 SAEs compared with and, 238/597 patients (39.9%) in the R-chemo+R arm experiencing 450 SAEs (see Table 4.10).

Table 4.10: Serious adverse events over the entire study period, occurring in ≥1% patients (safety population)

n, (%)	obin-chemo+obin n=595	R-chemo+R n=597
Total number of patients with at least one event	274 (46.1)	238 (39.9)
Total number of AE, n	590	450
Blood and lymphatic system disorders		
Number of patients with at least one AE	56 (9.4)	47 (7.9)
Febrile neutropenia	29 (4.9)	19 (3.2)
Neutropenia	22 (3.7)	25 (4.2)
Gastrointestinal disorders		
Number of patients with at least one AE	43 (7.2)	28 (4.7)
Diarrhoea	8 (1.3)	6 (1.0)
Abdominal pain	8 (1.3)	5 (0.8)
Vomiting	3 (0.5)	7 (1.2)
General disorders and administration site conditions		
Number of patients with at least one AE	30 (5.0)	34 (5.7)
Pyrexia	18 (3.0)	17 (2.8)

n, (%)	obin-chemo+obin n=595	R-chemo+R n=597
Infections and infestations		
Number of patients with at least one AE	108 (18.2)	86 (14.4)
Pneumonia	29 (4.9)	25 (4.2)
Herpes zoster	6 (1.0)	8 (1.3)
Urinary tract infection	8 (1.3)	5 (0.8)
Infection	5 (0.8)	7 (1.2)
Lower respiratory tract infection	8 (1.3)	3 (0.5)
Lung infection	5 (0.8)	6 (1.0)
Sepsis	8 (1.3)	2 (0.3)
Bronchitis	6 (1.0)	3 (0.5)
Gastroenteritis	7 (1.2)	1 (0.2)
Injury, poisoning and procedural complications		
Number of patients with at least one AE	41 (6.9)	21 (3.5)
Infusion-related reactions	27 (4.5)	11 (1.8)
Respiratory, thoracic and mediastinal disorders		
Number of patients with at least one AE	33 (5.5)	30 (5.0)
Dyspnoea	6 (1.0)	6 (1.0)
Pulmonary embolism	6 (1.0)	2 (0.3)
Vascular disorders		
Number of patients with at least one AE	12 (2.0)	7 (1.2)
Hypotension	6 (1.0)	0

The frequency and severity of adverse events of particular or special interest in GALLIUM are reported in Table 4.11. Again, nearly all adverse events were more often reported in the obin-chemo+obin arm compared with the R-chemo+R arm.

Table 4.11: The frequency and severity of AE of particular or special interest in GALLIUM

n, (%)	obin-chemo+obin n=595	R-chemo+R n=597
Infusion-related reactions*		
Number of patients with at least one AE	406 (68.2)	349 (58.5)
Number of patients with Grade 3–5 AEs	40 (6.7)	22 (3.7)
Number of patients with serious AEs	33 (5.5)	0 (0.0)
Neutropenia		
Number of patients with at least one AE	301 (50.6)	269 (45.1)
Number of patients with Grade 3–5 AEs	261 (43.9)	226 (37.9)
Number of patients with serious AEs	50 (8.4)	44 (7.4)
Infections		
Number of patients with at least one AE	460 (77.3)	418 (70.0)
Number of patients with Grade 3–5 AEs	118 (19.8)	93 (15.6)
Number of patients with serious AEs	108 (18.2)	86 (14.4)
Tumour lysis syndrome		
Number of patients with at least one AE	6 (1.0)	3 (0.5)
Number of patients with Grade 3–5 AEs	6 (1.0)	3 (0.5)

n, (%)	obin-chemo+obin n=595	R-chemo+R n=597
Number of patients with serious AEs	3 (0.5)	1 (0.2)
Thrombocytopenia		
Number of patients with at least one AE	68 (11.4)	45 (7.5)
Number of patients with Grade 3–5 AEs	36 (6.1)	16 (2.7)
Number of patients with serious AEs	4 (0.7)	1 (0.2)
Acute thrombocytopenia		
Number of patients with at least one AE	7 (1.2)	0 (0.0)
Number of patients with Grade 3–5 AEs	5 (0.8)	0 (0.0)
Number of patients with serious AEs	2 (0.3)	0 (0.0)
Haemorrhagic events		
Number of patients with at least one AE	57 (9.6)	62 (10.4)
Number of patients with Grade 3–5 AEs	5 (0.8)	7 (1.2)
Number of patients with serious AEs	6 (1.0)	5 (0.8)
Gastrointestinal perforation		
Number of patients with at least one AE	4 (0.7)	3 (0.5)
Number of patients with Grade 3–5 AEs	3 (0.5)	0 (0.0)
Number of patients with serious AEs	3 (0.5)	0 (0.0)
Cardiac events		
Number of patients with at least one AE	78 (13.1)	58 (9.7)
Number of patients with Grade 3–5 AEs	22 (3.7)	17 (2.8)
Number of patients with serious AEs	0 (0.0)	0 (0.0)
Second malignancies (6 months after first study drug intake)		
Number of patients with at least one AE	62 (10.4)	42 (7.0)
Number of patients with Grade 3–5 AEs	30 (5.0)	17 (2.8)
Number of patients with serious AEs	35 (5.7)	18 (3.0)
Hepatitis B reactivation		
Number of patients with at least one AE	3 (0.5)	2 (0.3)
Number of patients with Grade 3–5 AEs	0 (0.0)	0 (0.0)
Number of patients with serious AEs	0 (0.0)	0 (0.0)

*Most frequent symptoms of IRRs; nausea (24.2% [obin-chemo+obin], 19.3% [R-chemo+R]), chills (15.0%, 6.9%), pyrexia (13.6%, 5.5%), vomiting (10.4%, 7.5%), fatigue (6.7%, 6.9%)

The overall safety profile by chemotherapy subgroup is provided in Tables 54 and 55 of the CS. Treatment with bendamustine was associated with a higher incidence of Grade 3–5 infections and second malignancies during the maintenance and follow-up phases, while CHOP regimens were associated with higher rates of Grade 3–5 neutropenia during induction. Furthermore, non-relapse fatal AEs were more common in bendamustine treated patients (Obin-benda 5.9% vs. R-benda 4.4%) than in those treated with CHOP (1.6% vs. 2.0%) or CVP (1.6% vs. 1.8%).

The European Medicines Agency (EMA) requested further safety analyses, which resulted in a 'Revised Safety Analysis', which is an analysis conducted on the safety data derived from a 5 May 2017 snapshot. These updated results are presented in Table 4.12.

Table 4.12: Overview of adverse events in patients with Follicular Lymphoma in the GALLIUM trial (FL safety population)

Snapshot Date	**29 April 2016**				**05 May 2017**			
Safety parameters	**Obin-chemo n = 595**		**R-chemo n = 597**		**Obin-chemo n = 595**		**R-chemo n = 597**	
No. of AEs	10311		9341		11100 (+789 events; +7.7% rel. change)		10081 (+740 events; +7.9% rel. change)	
No. of patients with at least 1(%):								
AE (all grades)	592 (99.5%)		585 (98.0%)		594 (99.8%) (+0.3%)		592 (99.2%) (+1.2%)	
Grade 3-5 AE	444 (74.6%)		405 (67.8%)		456 (76.6%) (+2.0%)		418 (70.0%) (+2.2%)	
Fatal AE	24 (4.0%)		20 (3.4%)		24 (4.0%) (+0%)		20 (3.4%) (+0%)	
SAE	274 (46.1%)		238 (39.9%)		277 (46.6%) (+0.5%)		239 (40.0%) (+0.1%)	
AE leading to any withdrawal from any treatment	97 (16.3%)		85 (14.2%)		95 (16.0%) (+0.3%)		86 (14.4%) (+0.2%)	
AEs of particular interest, n (%):	**All Grades**	**Grade ≥3**	**All Grades**	**Grade ≥3**	**All Grades**	**Grade ≥3**	**All Grades**	**Grade ≥3**
Infusion-related reaction	406 (68.2%)	74 (12.4%)	349 (58.5%)	40 (6.7%)	420 (70.6%) (+2.4%)	73 (12.3%) (+0.2%)	361 (60.5%) (+2.0%)	44 (7.4%) (+0.7%)
Neutropenia	301 (50.6%)	273 (45.9%)	269 (45.1%)	236 (39.5%)	311 (52.3%) (+1.7%)	284 (47.7%) (+1.8%)	281 (47.1%) (+2.0%)	247 (41.4%) (+1.8%)
Infection	460 (77.3%)	119 (20.0%)	418 (70.0%)	93 (15.6%)	477 (80.2%) (+2.9%)	123 (20.7%) (+0.7%)	435 (72.9%) (+2.8%)	98 (16.4%) (+0.8%)
TLS	6 (1.0%)	6 (1.0%)	3 (0.5%)	3 (0.5%)	6 (1.0%) (+0%)	6 (1.0%) (+0%)	3 (0.5%) (+0%)	3 (0.5%) (+0%)
Thrombocytopenia	68 (11.4%)	36 (6.1%)	45 (7.5%)	16 (2.7%)	74 (12.4%) (+1.0%)	36 (6.1%) (+0%)	48 (8.0%) (+0.5%)	17 (2.8%) (+0.2%)
Acute thrombocytopenia	7 (1.2%)	5 (0.8%)	0	0	7 (1.2%) (+0%)	4 (0.7%) (+0.2%)	0	0
Hemorrhagic events	57 (9.6%)	5 (0.8%)	62 (10.7%)	7 (1.2%)	68 (11.4%) (+1.8%)	6 (1.0%) (+0.2%)	68 (11.4%) (+1.0%)	7 (1.2%) (+0%)
GI perforation	4 (0.7%)	3 (0.5%)	3 (0.5%)	0	5 (0.8%) (+0.2%)	3 (0.5%) (+0%)	3 (0.5%) (+0%)	0
Cardiac events (incl. IRRs)	78 (13.1%)	22 (3.7%)	58 (9.7%)	17 (2.8%)	85 (14.3%) (+1.2%)	24 (4.0%) (+0.2%)	60 (10.1%) (+0.4%)	17 (2.8%) (+0%)
(excl. IRRs)	57 (9.6%)	18 (3.0%)	49 (8.2%)	15 (2.5%)	63 (10.6%) (+1.0%)	20 (3.4%) (+0.3%)	51 (8.5%) (+0.3%)	15 (2.5%) (+0%)
Second malignancy (SOC)[a]	62 (10.4%)	30 (5.0%)	42 (7.0%)	17 (2.8%)	66 (11.1%) (+0.7%)	30 (5.0%) (+0.7%)	45 (7.5%) (+0.5%)	17 (2.8%) (+0%)
(SMQ)[a]	43 (7.2%)	28 (4.7%)	30 (5.0%)	16 (2.7%)	45 (7.6%) (+0.4%)	28 (4.7%) (+0%)	31 (5.2%) (+0.2%)	16 (2.7%) (+0%)

Source: EMA-responses-assessment-report by Roche.[21]

AE = adverse event; FL = follicular lymphoma; GI = gastrointestinal; IRR = infusion-related reaction; SAE = serious adverse event; SMQ = standardized MedDRA query; SOC = system organ class; TLS = tumour lysis syndrome.

Notes: Percentages in parentheses refer to the proportion of patients with at least 1 AE. For the 05 May 2017 snapshot, the change in the proportion within the treatment arm is also provided in a second parenthesis. Safety data from the 29 April 2016 snapshot were coded using MedDRA v18.1; safety data from the 05 May 2017 snapshot were coded using MedDRA v20.0.

[a] AEs reported under the SOC "Neoplasms benign, malignant and unspecified (incl. cysts and polyps)" starting at least 6 months after the first dose of study drug (which included both malignant and benign tumours) and AEs reported under the SMQ "Tumours malignant and unspecified", starting at least 6 months after the first dose of study drug are shown.

ERG comment:

- The ERG draws to the attention of the committee that although overall rates of adverse events between groups were similar, a higher rate of serious adverse events was noted with obinutuzumab (46.1% vs 39.9%). These led to a higher rate of dose withdrawal, reduction or interruption in the obinutuzumab group. Grade 3 to 5 AEs were also more frequent with obinutuzumab (74.6% vs 67.8%).
- Infusion-related events were more common with obinutuzumab (68.2% vs 58.5%). Other events occurring more frequently with obinutuzumab included neutropaenia, thrombocytopaenia and febrile neutropaenia.
- Although, overall there were fewer deaths with obinutuzumab, fatal AEs were slightly higher (24 (4%) vs 20 (3.4%).
- The committee will need to consider whether the results observed would affect management of FL and the importance of the adverse event profile to patients.

4.2.3 Overview of the non-randomised evidence

One non-randomised trial (GAUDI) was presented in the CS. The methodology of the open label trial is presented in Table 4.13 and the results are given in Table 4.14.

Table 4.13: Methodology of the non-randomised trial GAUDI

PICOS	Details
Population	81 previously untreated patients with FL
Intervention (1)	Obin-chemo (bendamustine) as induction followed by obin maintenance monotherapy (n = 41)
Intervention (2)	Obin-chemo (CHOP) as induction followed by obin maintenance monotherapy (n = 40)
Outcomes	• Safety of induction treatment (primary outcome) • Overall response rate • Complete response rate • Progression-free survival • Progression / death • Deaths due to progressive disease • Pharmacokinetics • B-cell depletion and recovery • Safety of maintenance treatment
Study design	Open-label non-randomised Phase 1b study
Source: CS[1] and Grigg 2017[22]	

The aim of the trial was to investigate obin-chemo with CHOP or bendamustine as induction followed by obin maintenance monotherapy. Assignment to chemotherapy centre was decided on a per centre basis before enrolment.

The patient profile was similar to GALLIUM. Ninety-one percent of patients had Ann Arbor stage III to IV, 67% had extra-nodal involvement and 43% had bulky disease, 82% had an intermediate/high FLIPI score.

Table 4.14: Results of the non-randomised trial GAUDI

Endpoint	Obin-benda (n = 41)	Obin-CHOP (n = 40)	Total (n = 81)
Efficacy			
ORR (%) (95% CI)	93 (80.1 to 98.5)	95 (83.1 to 99.4)	94 (86.2 to 98.0)
CR at end of induction, (%) (95% CI)	37 (22.1 to 53.1)	35 (20.6 to 51.7)	36 (25.4 to 47.2)
CR at 30 months, (%) (95% CI)	63 (46.0 to 78.2)	58 (40.8 to 74.5)	61 (NA to NA)
PFS at 36 months, % (95% CI)	90 (0.80 to 0.99)	84 (0.72 to 0.96)	87 (0.79 to 0.94)
Progression / death (n)	6	11	17
Deaths due to PD (n)	1	2	3
Safety			
Induction Grade 3 / 4 AE n (%)	21 (51)	31 (78)	52 (64)
Maintenance Grade 3 -5 AE n (%)	NR	NR	27 of 72 (37.5)

Source: Table 44 of CS
AE = adverse event; CR = complete response; ORR = overall response rate; PD = progressive disease; PFS = progression-free survival

All patients experienced an adverse event in the induction phase. 64% experienced grade 3/4 events. The most common adverse event during induction was infusion-related infections (58%). Fifty patients had 74 dose delays or interruptions of obinutuzumab due to adverse events (no dose reductions were allowed). The most common grade 3 haematological adverse event was neutropenia. During maintenance 27 of 72 patients experienced grade 3-5 AEs, with nine withdrawing from treatment due to an AE. Eight patients had haematological events during maintenance (all obin-benda group). The CS concluded that induction therapy with obin-benda or obin-CHOP followed by obin maintenance was associated with tolerable safety.

ERG comment:

- The non-randomised trial does not add considerably to the information in the submission as it is small, non-randomised and cannot be used meaningfully to compare chemotherapy regimens. However it can be noted that the AE profile and overall response at induction are similar to that observed in GALLIUM.

4.2.4 Ongoing trials

The GALLIUM trial is ongoing. The CS stated that 'Further analysis from an updated data cut (clinical cut-off 10 September 2016) that formed the basis of the economic analysis will be available within the next 12 months, as well as a 90-day safety update for the FDA.'[1] We have presented in this report any data we have received from the company in relation to the 10 September cut-off. The company was also asked if any further data were available. The company responded that there are no data from later data-cuts from the GALLIUM study available at this point in time.[8] However, [redacted].

The CS stated that there were no further studies investigating obinutuzumab in the indication under appraisal.

4.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

Only one trial is included in the CS: the GALLIUM trial. No indirect comparisons and/or multiple treatment comparisons were performed.

4.4 Critique of the indirect comparison and/or multiple treatment comparison

Only one trial is included in the CS: the GALLIUM trial. No indirect comparisons and/or multiple treatment comparisons were performed.

4.5 Additional work on clinical effectiveness undertaken by the ERG

No further additional work was undertaken by the ERG.

4.6 Conclusions of the clinical effectiveness section

The CS includes a systematic review of the RCT evidence for obinutuzumab in previously untreated FL. One RCT was identified (GALLIUM) that investigated the efficacy and safety of obinutuzumab with chemotherapy followed by obinutuzumab maintenance for responders. This was a large, well-conducted randomised trial including 1,202 patients with FL. The trial was conducted at 177 centres in 18 countries (293 (21%) patients were from the UK and almost 50% were from Western Europe).

A number of limitations were identified by the ERG. The GALLIUM trial was open-label, therefore results based on independent review will be less prone to bias than investigator results. In the trial, obinutuzumab and its comparator rituximab could be given to patients with three different chemotherapy regimens (CHOP, CVP and bendamustine). In the trial approximately 57% received bendamustine, 33% CHOP and 10% CVP. The breakdown of the chemotherapy used may not be reflective of the UK. The trial was not designed to investigate differences in chemotherapy regimens so any variation in results between chemotherapy regimens may reflect genuine differences of effectiveness or patient selection factors.

Although GALLIUM had a reasonable follow up duration, data were not fully mature for the main outcomes. Median progression-free survival (PFS) could not be determined. Overall obinutuzumab was superior to rituximab for PFS (HR = 0.72 (0.56 to 0.93)) for the latest cut-off using IRC data. Although outcomes relating to progression were positive, no differential effects on HRQoL were identified. The committee will need to consider whether improvements in PFS and possible delay to new anti-lymphoma medication are worthwhile alone. The committee will further need to consider any possible relationship between improvements in PFS and improvements in overall survival. Overall survival data in GALLIUM were not mature. GALLIUM is an ongoing trial which should provide, further, more mature results. Finally, the higher rate of serious and higher grade events with obinutuzumab needs to be considered in terms of management of the disease and acceptability to patients.

5. COST EFFECTIVENESS

5.1 ERG comment on company's review of cost effectiveness evidence

This section pertains mainly to the review of cost effectiveness analysis studies. However, the search section (5.1.1) also contains summaries and critiques of other searches related to cost effectiveness presented in the company submission. Therefore, the following section includes searches for the cost effectiveness analysis review, measurement and evaluation of health effects as well as for cost and healthcare resource identification, measurement and valuation.

5.1.1 Searches performed for cost effectiveness section

The following paragraphs contain summaries and critiques of all searches related to cost effectiveness presented in the company submission.

Searches for cost effectiveness analysis review

A literature review was conducted to identify all published studies that assessed the cost effectiveness of treatments in first-line follicular lymphoma.

Searches were reported for PubMed, Embase, and the NHS Economic Evaluation Database (NHS EED). The host provider for each database and the date the searching was conducted was provided in the clarification response.[8, 13] Database date spans were not reported. Conflicting search dates were noted between the original searches in the CS Appendix[12] and the clarification response.[13]

In order to address the limitations queried by the ERG in the clarification process, further revised searches were carried out in June 2017. The date span for each database was not reported. The revised PubMed search strategy was limited to 1998 onwards. The revised Embase strategy included a different date limit, which may have been a typographical error ([1-1-2017]/sd NOT [7-3-2017]/sd), as the results for this search line did not appear to match up to a single month's database references. The NHS EED strategy did not include a date restriction.

In the original submission, the Embase strategy was duplicated, and was erroneously reported as having been conducted in the NHS Economic Evaluation Database (NHS EED) search. Following clarification, the company stated the wrong strategy had been reported, and provided a different NHS EED search strategy.[8, 13]

Supplementary searches were carried out in five conference proceedings and a number of relevant organisational websites. No date of search was reported for the supplementary searches, and details of search terms were provided in response to the clarification process.[8, 10]

These meet the requirements detailed in the NICE guide to the methods of technology appraisal.[11]

Search strategies for the database searches were provided in the Appendix 3 of the CS[12] and in two Excel spreadsheets supplied as part of the clarification response.[8, 10, 13]

For the most part, the searches were well reported and reproducible; the names of the database hosts were provided in the clarification response.[8] The database searches were well structured and used combinations of index terms appropriate to the resource searched, free text and a number of synonyms for the condition and terms from an economics/cost study design filter[23] were used to further restrict the search results.

The disease terms used for these PubMed and Embase searches were the same as those employed in the clinical effectiveness search, therefore the same observations apply.

The cost effectiveness searches were restricted to references with line of treatment (newly diagnosed or untreated patients) in the title or abstract. The ERG remained concerned that restriction of all searches to line of treatment was overly restrictive and problematic. There were no appropriate indexing terms for this concept in PubMed or Embase, therefore this restriction was entirely dependent on free-text terms.

The revised NHS EED search undertaken to inform the cost effectiveness review was different to the strategy employed for the clinical effectiveness CENTRAL search, however similar significant errors in the use of search syntax and within-phrase wildcard use were observed.

Lines #2, #3, #4 and #9 all utilise the wildcard command "?" or another truncation symbol"*" within phrase terms. The Cochrane Library search help clearly states "Phrase search does NOT support the use of wildcards".[18] This problem affects the disease facet, as well as the lines restricting the search to untreated or newly diagnosed disease. The ERG did not consider the NHS EED search adequately robust to inform the cost effectiveness systematic review, as the only unaffected lines involve subject indexing. The ERG noted a typographical error in the line combination on line #10 of the NHS EED search.

#10	((#1 OR #2) AND #3) OR #4 to #7	294	358

The ERG assumed this line should read:

((#1 or #2) and #3) **or #4 or #5 or #6 or #7**

It is unclear whether this was a reporting error or whether the search had been executed using this command. If used as part of the search strategy, this would be a consequential error, potentially affecting the retrieval of relevant records.

Measurement and valuation of health effects

A search to identify health-related quality of life (HRQoL) studies relevant to the decision problem was conducted. Searches were reported for PubMed, Embase and the Cochrane Library. The section of the Cochrane Library used was not reported, however based on the numbers presented, the ERG assumed NHS EED was searched. Six conference proceedings were searched, and the terms used were reported in the clarification response.[10] The clarification response reported the host and search dates for all databases. The searches were well reported and reproducible.

The database searches were well structured and used combinations of index terms appropriate to the resource searched, free text and a number of synonyms for the condition and terms based on an HRQoL/Health state utilities filter[24] was used to further restrict the search results.

As noted in the clinical and cost effectiveness searches of CENTRAL and NHS EED, significant problems resulting from incorrect use of within-phrase wildcard characters were also noted in the revised Cochrane/NHS EED search, presented in the clarification response.[13] Lines #2, #3 and #4 all utilise the wildcard command "?" within phrase terms. The Cochrane Library search help clearly states "Phrase search does NOT support the use of wildcards".[18] This problem affects the disease facet, however the utilities/HRQoL facet is unaffected. The ERG did not consider the Cochrane/NHS EED search adequate.

Cost and healthcare resource identification, measurement and valuation

Resource use searches were presented in the CS appendix document,[12] and were not revised as part of the clarification process. These searches were performed in March 2017 on Embase, Medline, NHS EED and EconLit. Date spans and database hosts were not reported. Searches of Embase and Medline were limited to 1998 onwards and to English language publications only. The NHS EED search was limited by date to 1998 onwards, and the EconLit did not include language or date restrictions. From the database syntax used, Medline and Embase appeared to have been searched via the Ovid host. The searches were well reported and reproducible. The database searches were clearly structured and used combinations of index terms and free text.

As noted with the clinical and cost effectiveness searches, this search was restricted to references with line of treatment (newly diagnosed or untreated patients) in the title or abstract. The ERG considered this overly restrictive as it is possible that a relevant study might not describe line of treatment in the title or abstract. The ERG raised this point during clarification, and the company disagreed and felt RCTs in follicular lymphoma would state line of treatment in the title or abstract. The company thought it was unlikely relevant resource use studies would be missed by restricting the search to line of treatment, as their results overlapped with a separate submission, which was also restricted in this way. The ERG remained concerned that restriction of all searches to line of treatment was overly restrictive and problematic. There were no appropriate indexing terms for this concept in PubMed or Embase, therefore this restriction was entirely dependent on free-text terms.

In addition the ERG noted this facet also showed significant problems resulting from incorrect use of within-phrase wildcard characters. Line #6 utilised the wildcard command "?" without spacing within phrase terms. The ERG queried this issue during the clarification process, and the company responded that they had been unable to reproduce the issue in Ovid and that they had validated the use of wild card characters for the resource use review.[8] The ERG was unclear what method of validation was used and this information was not supplied in the clarification response, nevertheless the errors were still present and impacted on the retrieval of all the search strategies. Incorrect use of the wildcard within a phrase without spacing has resulted in a consequential error in the Medline, Embase, EconLit and NHS EED searches.

In the example presented below, the first row presents the company's search line repeated in Embase (date of the ERG's search: 10 July 2017). The errors in use of wildcards and spacing are highlighted. The second row presents the ERG's corrections to wildcard use to demonstrate the potential differences in numbers retrieved by the lines. It was not possible for the ERG to correct, repeat and re-screen all the company's searches in the time available, however the ERG considered this a consequential error undermining the robustness and rigour of these searches. As this search line was reproduced in the Medline, Embase, EconLit and NHS EED strategies, all the resource use searches were affected.

3	(untreat or first?line or na?ve or "not treated" or not?treated or "not exposed" or unexposed or "new?diagnos*" or "de?novo" or "newly diagnosed" or primary or initial or early or "never?treated" or frontline or front-line or "front line" or ("without prior" and (regimen or therap* or treatment*))).ti,ab.	4036979
4	(untreat or first-line or na*ve or "not treated" or not-treated or "not exposed" or unexposed or "new* diagnos*" or "de?novo" or "newly diagnosed" or primary or initial or early or "never* treated" or frontline or front-line or "front line" or ("without prior" and (regimen or therap* or treatment*))).ti,ab.	4308282

The errors in wildcard use compound the restriction of the line of treatment facet.

Another restriction raised by the ERG during the clarification process was the application of an English language restriction to the Resource Use Medline and Embase search strategies presented in Table 16 of the original CS Appendices.[12] Current best practice states that '*Whenever possible review authors should attempt to identify and assess for eligibility all possibly relevant reports of trials irrespective of*

language of publication'.[25] In the response to clarification, the company responded that their intention was to identify UK studies only and that it was unlikely UK based studies would report in languages other than English. To check the consequences of this language restriction, the company reported re-running the search without the language restriction and rescreening the missed non-English references. The clarification response reported that none of the non-English references were relevant to the research question. Whilst that check was reassuring, the ERG felt that it would have been preferable to prospectively minimise the introduction of potential language bias when running the searches, rather than checking whether bias was observable in this instance. Application of the English language restriction to the Embase and Medline strategies only removed 32 and 26 references respectively, which would not have been too onerous a task for the company to screen when the searches were originally conducted.

5.1.2 Inclusion/exclusion criteria used in the study selection

In Table 5.1 inclusion and exclusion criteria for the study selection of the cost effectiveness review are presented.

Table 5.1: Inclusion and exclusion criteria used for the cost effectiveness review

PICOS	Inclusion criteria	Exclusion criteria
Population	People in the United Kingdom with iNHL who were previously untreated. All subtypes, except skin lymphomas	Disease area not iNHL Relapsed or refractory setting Setting not UK
Intervention & Comparators	Intervention and comparator not restricted	
Outcomes	The outcome measures to be considered for the economic literature review are: Costs Resource use Quality of life Utility	
Study types	Health economic evaluations	Other study types: Secondary publications Review articles, systematic literature reviews, or meta-analyses Editorials, notes or letters to the editor Studies containing no primary data
Source: Based on Table 58 of the CS[1] iNHL = indolent non-Hodgkin lymphoma		

ERG comment: The inclusion and exclusion criteria used by the company seem to be appropriate for the selection of the cost effectiveness studies.

5.1.3 Included/excluded studies in the cost effectiveness review

With its search, the company identified a total of 1,861 records, two of which were identified by sources other than the electronic databases. After exclusion, 42 full-text publications were assessed for their eligibility. Finally, six articles were included in the narrative summary. The CS states that none of the identified studies met the inclusion criteria since no study investigated the cost effectiveness of obin-

chemo+obin as an intervention in previously untreated patients with FL.[1] However, a list of six UK-based cost effectiveness studies in previously untreated patients with FL were included in the CS.[1]

Dundar et al.,[26] Ray et al.,[27] and Papaioannou et al.[28] studied the cost effectiveness of rituximab in combination with various chemotherapy regimens as induction. While the former two used a Markov model for their analysis, Papaioannou et al.[28] made use of a patient level simulation.

Using a Markov model, Greenhalgh et al.[29] estimated the cost effectiveness of rituximab as a maintenance therapy. Dewilde et al.[30] studied rituximab as induction therapy, followed by maintenance with a patient level simulation approach. Finally, Wang et al.[31] derived both costs and outcomes through a patient level simulation from a UK observational cohort.

All studies were assessed for their quality with the Drummond and Jefferson (1996) checklist[32] and are reported in Appendix 5 of the CS.[12]

5.1.4 Conclusions of the cost effectiveness review

The cost effectiveness searches in the CS were well documented and reproducible, and were carried out in line with the NICE guide to the methods of technology appraisal.[11] Issues resulting from incorrect use of within-phrase wildcard characters and restriction of the cost effectiveness search to line of treatment may have impaired the search recall. Language limits in the resource use searches may have led to relevant evidence being unidentified.

Besides a summary table of the six identified studies (Table 59 in the CS[1]) and a quality assessment of these studies in the CS appendix,[12] no specific conclusions of the cost effectiveness review were provided. Therefore, the company developed a de novo economic model to address the decision problem.

5.2 Summary and critique of company's submitted economic evaluation by the ERG

A summary of the de novo economic model developed by the company is presented in Table 5.2.

Table 5.2: Summary of the company submission economic evaluation

	Approach	Source/Company's justification	Signpost (location in CS)
Model	The model developed for this submission was a four-state cohort transition (Markov) model with monthly cycles. Time horizon in the base-case was 40 years. The average age of the cohort was 57.9 years. Baseline patient characteristics were taken from the GALLIUM trial.[33]		Section 5.2. (p. 134)
States and events	The health states included in the model are progression free (PFS) (on/off treatment), two progressed disease (PD) states, early PD and late PD, and death. All patients begin in the PFS health state on treatment and are assigned to a PFS 'on-treatment' utility value and treatment costs while on therapy. Time to treatment discontinuation is based on the actual observation from the GALLIUM study for both arms. [33]Specifically, as per license indication, only patients responding to induction received maintenance. Maintenance was only offered until progression or for a maximum of two years; then it is said that treatment is completed. When patients complete or discontinue treatment in the PFS state, they are considered off treatment and assigned an 'off treatment' PFS utility value and costs for ongoing monitoring in supportive care. Patients can either remain in PFS (on- or off-treatment) or exit the state due to disease progression or death. Two progressed disease states were introduced to account for different outcomes and costs to the cohorts of patients who experience an early or a late progression. Once patients enter any of the two PD states, patients can only remain in their corresponding PD state until death.	The model structure is in line with a typical oncology Markov model. The model developed in this CS incorporates early and late PD. This distinction is made because time to progression is highly predictive for post progression mortality and overall survival. Patients progressing within two years of initial treatment, have significantly worse mortality rates than patients who did not progress within two years. [3 34]	Section 5.2 (p. 135)
Comparators	Rituximab-based chemotherapy, with rituximab maintenance treatment	Although more comparators are considered in the NICE scope, [35] the company presented evidence for one of them: Rituximab-based chemotherapy, with rituximab maintenance treatment (which is the comparator in the GALLIUM trial).	Section 5.2 (p. 137)

	Approach	Source/Company's justification	Signpost (location in CS)
Natural History	Advanced stage FL is a progressive condition. Patients are usually considered incurable and therefore standard therapeutic approaches attempt to control the condition. Advanced stage FL is typified by a chronic course of repeated relapses, treatment and progression. Median life expectancy ranges from 8–12 years after diagnosis, although this has extended to around 15 years in the post-rituximab era.[4] The 5-year survival rate of patients with FL in the UK is estimated to be 87.2%.[5]		Section 3.1 (p. 35)
Treatment effectiveness	Results demonstrated superior PFS with obin-chemo compared with R-chemo. This is the main driver of the differences in costs and QALYs between the treatment arms.	PFS probabilities were predicted based on regression analyses with treatment effect as a covariate, performed on data from the GALLIUM trial.	Section 5.3 (p. 138)
Adverse events	The model includes the most common adverse events observed in the GALLIUM trial. The cost and disutility effects (the latter only in scenario analysis) of these adverse events were incorporated in the PFS (on-treatment) health state for a maximum of two years.	All adverse events of Grades 3, 4, or 5 occurring in more than 2% of patients in either arm of the GALLIUM trial were incorporated into the model. Justification of the choice of the 2% as cut-off value was not provided.	Section 5.4 (pp. 162 and 180)
Health related QoL	The model uses EQ-5D utilities collected from the GALLIUM trial for the PFS health state. PD health states utilities and adverse event disutilities were sourced from the literature.	Long-term EQ-5D utility scores collected in the GALLIUM trial were considered immature. Therefore, post progression utilities were sourced from Wild et al.[36] This study was deemed appropriate by the company.	Section 5.4 (p. 151)
Resource utilisation and costs	Health state related costs consisted of medication costs (induction and maintenance), supportive care costs, subsequent treatment costs in PD, and adverse event costs. Relevant medication costs included costs of obinutuzumab, bendamustine, CHOP, CVP, and rituximab.	Based on UK reference costs	Section 5.5 (p. 165)
Discount rates	A 3.5% discount rate was used for both utilities and costs.	According to NICE reference case	Section 5.2 (p. 137)
Sensitivity analysis	One-way deterministic sensitivity analysis, scenario analyses and probabilistic sensitivity analysis.	Ranges/scenarios based on different assumptions.	Section 5.8 (p. 192)

Source: table derived from the CS. [1]

EQ-5D = European Quality of Life-5 Dimensions; NICE = The National Institute for Health and Care Excellence; QoL = Quality of Life; TA = Technology Appraisal.

5.2.1 NICE reference case checklist (TABLE ONLY)

Table 5.3: Comparison of the CS model with the NICE reference case

Elements of the economic evaluation	Reference Case	Included in submission	Comment on whether de novo evaluation meets requirements of NICE reference case
Comparator(s)	Therapies routinely used in the NHS, including technologies regarded as current best practice	Partly	Only rituximab-based chemotherapy, with rituximab maintenance treatment was considered as a comparator. More comparators were considered in the NICE scope.
Type of economic evaluation	Cost effectiveness analysis	Yes	
Perspective on costs	NHS and PSS	Yes	
Perspective on outcomes	All health effects on individuals	Yes	
Time horizon	Sufficient to capture differences in costs and outcomes	Partly/Yes	The time horizon considered was 40 years. However, at the end of the time horizon between 3% and 5% of the patients were still alive in the model. This was deemed high by the ERG. After the clarification phase, the time horizon of the model was increased to 50 years.
Synthesis of evidence in outcomes	Systematic review	Yes	Meta-analysis was not used; all effectiveness data used in the model were based on two trials: GALLIUM and PRIMA.
Measure of health effects	QALYs Life-years	Yes	
Source of data for measurement HRQOL	Reported directly by patients and/or carers.	Yes/Unclear	PFS utility data were based on EQ-5D utilities collected from the GALLIUM trial (N=1097). PD health states utilities were based on EQ-5D collected data on 222 patients with FL in eight centres in the UK.[36]
Source of preference data for valuation of changes in HRQOL	Sample of public	Unclear	It was not clear which data were used in the valuation of the EQ-5D.
Discount rate	Annual rate of 3.5% on costs and utilities	Yes	
Equity weighting	No special weighting	Yes	
Sensitivity analysis	Probabilistic sensitivity analysis	Yes	Additionally, univariate sensitivity and scenario analyses were performed.

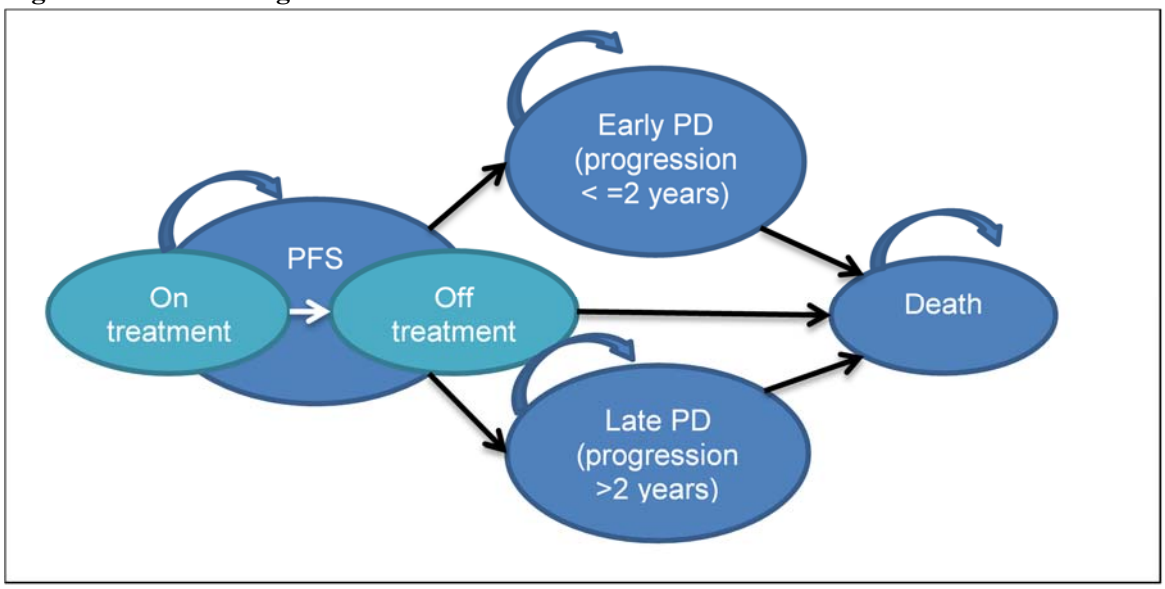
EQ-5D = European Quality of Life-5 Dimensions; NHS = National Health Service; NICE = The National Institute for Health and Care Excellence; PSS = Personal Social Services; QALYs = Quality adjusted life years.

5.2.2 Model structure

The model developed for this submission was a four-state cohort transition (Markov) model. The health states included in the model are progression free (PFS) (on/off treatment), two progressed disease (PD) states, early PD and late PD (with subsequent treatments) and death. The model structure is shown in Figure 5.1. Initially all patients begin in the PFS health state on treatment (obin-chemo+obin or R-chemo+R) and are assigned a PFS 'on-treatment' utility value and treatment costs while on therapy. Time to treatment discontinuation is based on the actual observation from the GALLIUM study for both arms.[33] Specifically, as per license indication, only patients responding to induction received maintenance. Maintenance was only offered until progression or for a maximum of two years; then it is said that treatment is completed. When patients complete or discontinue treatment in the PFS state, they are considered off treatment and assigned an 'off treatment' PFS utility value and costs for ongoing monitoring in supportive care. Patients can either remain in PFS (on- or off-treatment) or exit the state due to disease progression or death.

Time to progression after initial treatment is highly predictive for post progression mortality and overall survival. In particular, patients progressing early (within two years of initial treatment) have significantly worse mortality rates than patients who did not progress within two years.[3, 34] Therefore, two progressed disease states were introduced to account for different outcomes and costs to the cohorts of patients who experience an early or a late progression. Once patients enter any of the two PD states, patients can only remain in their corresponding PD state until death. The death state is an absorbing state meaning that the proportion of patients in this state is calculated by the sum of deaths in the PFS and PD states. The cumulative deaths from PFS, early and late PD states are used to calculate overall survival in the model.

Figure 5.1: Model diagram



PD = Progressed Disease; PFS = Progression free Survival. Source: Figure 22 in the CS. [1]

ERG comment: The model structure in the CS is in line with other, commonly used, Markov models for progression in oncology. However, other models usually consider three health states: PFS, PD and death. The model developed in the CS has the advantage of incorporating early and late PD, which seems to be appropriate for the decision problem at hand.

5.2.3 Population

The patient population considered in the company's de novo economic analysis is the same population as in the GALLIUM trial, i.e. people with untreated advanced follicular lymphoma except for FL grade 3b. Apart from the exclusion of FL grade 3b, the population is in line with the scope. The baseline characteristics used in the health economic model are summarised in Table 5.4.

Table 5.4: Baseline characteristics

Patient characteristic	Baseline value	Source
Age (years)	57.9	GALLIUM trial[33]
Body weight (kg)	75.7	
Height (cm)	168.3	
Calculated Body Surface Area (m^2)	1.86	
Source: Table 60 in the CS.[1]		

ERG comment: The proportion of UK patients in the GALLIUM trial (21%) seems reasonable and nearly half of the patients are from Western Europe. However, based on the figures reported in the CS, the ERG considers that there might be differences between the GALLIUM population and the advanced FL population in the UK, which are worth exploring in the cost effectiveness analyses. For instance, page 77 of the CS states that the median age in the GALLIUM trial is 59 years but page 32 states that the median age at diagnosis in the UK is 65 years.[1] In the clarification response,[8] the company explained that the latter median is based on HMRN data, which relates to all FL patients at diagnosis, irrespective of treatment or management of patients. Therefore, it also includes patients with less advanced disease that require no active treatment or patients that may only receive palliative care and not R-chemo. Nevertheless, the HMRN also reports patient's age and treatment for follicular lymphoma in the years 2004-2012.[37] The median age of patients treated with chemotherapy is reported to be 63.7 years and these patients may be more representative for advanced follicular lymphoma. The company acknowledged that the GALLIUM cohort might be younger than the average UK patient. This was also confirmed by some of the clinical experts consulted by the company (clarification response question A11).[8] Therefore, the ERG considers that a higher baseline age should have been used for the base-case analysis. This would result in a more conservative approach.

The company was not aware of literature reporting the other demographic variables included as parameters in the economic model, e.g. Body Surface Area (BSA). The company refers to a recent publication which reports BSA for patients treated for a range of cancers (but not haematological) in England as reported in the SACT database.[38] The average for women was 1.74m^2 and 1.95m^2 for men. Based on the proportion of 50.6% male patients in the GALLIUM trial, the UK average of 1.85m^2 derived from SACT is very similar to the 1.86m^2 in the GALLIUM trial. Although there seems to be an inconsistency with the proportion of males reported in the CS (46.8% of males on page 77 in the CS - which is in line with the incidence rates reported in Section 3.1 of the CS),[1] the ERG agrees with the company that it is unlikely that, based on these figures, the dosing of rituximab or chemotherapy would be significantly different in clinical practice compared to the GALLIUM trial. The ERG noticed that the BSA input parameter is not explicitly used in the model (i.e. the cell F30 in the model sheet "Model Inputs" is not linked to any model formula). Values for actual and planned dose in the model are hardcoded, meaning that they do not allow variation. Given that there seems to be some uncertainty with respect to this parameter, the ERG considers that this should have been included in the PSA. Due to lack of data and time constraints, the ERG could not implement this in the model. Therefore, it is likely that the current probabilistic results underestimate the overall uncertainty in the model.

The ERG considers that the proportions of patients treated in each chemotherapy regimen (bendamustine, CHOP, CVP) in the GALLIUM trial (57%, 33% and 10%, respectively – Table 25 of the CS) and in the general UK population (29%, 13% and 36%, respectively – Table 14 of the CS[1]) are quite different and might indicate that the proportions used in the GALLIUM trial are not reflective of the UK clinical practice. In the clarification response, the company explained that the proportions presented in Table 14 of the CS[1] were based on a questionnaire-based UK sample (Q4 2016 – Q1 2017, Haematology TAMS, Genactis [as cited in Clarification response]) of 157 cases reported by 45 clinicians. The company also indicated that in the GALLIUM trial, 68% of the UK patients were given bendamustine and 31% CVP, indicating a more preferential use of bendamustine compared to the market research sample. The company also mentioned that, according to the discussions in the advisory board, there are local variations in clinical practice with respect to chemotherapy use and therefore, the appropriate representative average use of the three chemotherapy regimens has some uncertainty. This implies that whether the proportions used in the GALLIUM trial are reflective or not of the UK clinical practice is also uncertain. Since GALLIUM was not powered to detect differences between the three chemotherapy methods and patients were not randomised to chemotherapies, the ERG considers that it is not feasible to conduct a robust scenario analysis where PFS and OS estimates are obtained with a different proportion of chemotherapy regimens. The ERG agrees with the company that the only feasible scenario analysis may be then to assume equal clinical outcomes while considering chemotherapy, administration and AE costs according to an alternative distribution of patients per chemotherapy group. However, if there is any treatment effect due to the underlying chemotherapy method, this would not be possible to detect with the current analyses. In the ERG preferred base-case, the proportion of UK patients in the GALLIUM trial was considered; thus, 68% bendamustine, 31% CVP and 1% CHOP. The main reason for this was that the GALLIUM study recruited 293 UK patients in 29 centres, which seems as representative for the UK as the market research questionnaire (which was considered in scenario analysis). Given that this proportion is uncertain, the ERG considers that these parameters should have been included in the PSA, for example as a Dirichlet distribution. Due to time constraints, the ERG could not implement this in the model. Therefore, the current probabilistic results are likely to underestimate the overall uncertainty in the model.

5.2.4 Interventions and comparators

The intervention included in the company's economic analysis was the same considered in the CS and in the GALLIUM trial: obinutuzumab in combination with chemotherapy (CVP, CHOP or bendamustine), followed by obinutuzumab maintenance therapy in patients achieving a response. The intervention is in line with the scope, where the intervention described is 'Obinutuzumab in combination with chemotherapy, with or without obinutuzumab maintenance therapy'. The comparator included in the company's economic analysis was rituximab-based chemotherapy, with rituximab maintenance treatment, which is the comparator considered in the GALLIUM trial. This comparator is in line with the scope.

ERG comment: As mentioned in Section 3.2, the company did not provide any evidence for obinutuzumab induction therapy without maintenance therapy. Therefore, this was not included in the economic analysis. Other relevant comparators listed in the NICE scope (obinutuzumab versus rituximab mono-therapy and bendamustine mono-therapy) were not included in the company's cost effectiveness analysis either.

5.2.5 Perspective, time horizon and discounting

The cost effectiveness analyses adopted the perspective of the NHS/PSS and a discount rate of 3.5% was applied for both costs and utilities. A 40-year time horizon was used.

ERG comment: At the end of the base-case simulation, 3.8% and 3.3% of the patients were still alive in the treatment and comparator arms, respectively. The ERG asked the company to adjust the model to perform the analysis with a longer time horizon. In the revised version of the model (after clarification), a 50-year time horizon was used.

5.2.6 Treatment effectiveness and extrapolation

Clinical parameters for the model were derived from the GALLIUM trial data when these were considered mature enough to provide robust estimates. Thus, GALLIUM data were used to estimate time to treatment discontinuation (TTTD), PFS and post progression survival (PPS) for early progressed disease. The investigator (INV) assessed PFS data (PFS-INV) was used, corresponding to the primary endpoint. The extrapolation beyond the observed period in the GALLIUM trial was based on parametric functions. The latest available data cut of GALLIUM with a clinical cut-off date of 10 September 2016 was used.

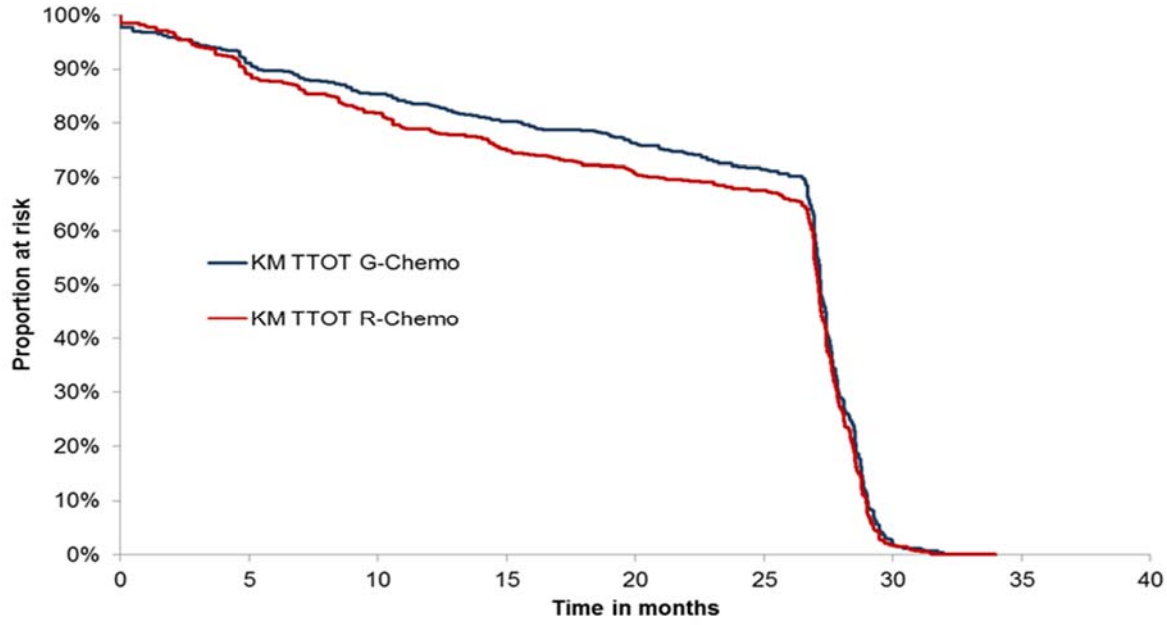
External data were used to populate PPS for late progressed disease using long-term data from the PRIMA trial.[39] To derive PPS for patients progressing late, data sources with longer follow up than GALLIUM were required to obtain sufficient death events for this group. Data from the PRIMA study were used in the base-case to estimate the mortality post progression for late PD as this data were based on a cohort receiving R maintenance after response to R-chemo induction treatment where patient level data with up to 9.75 years of follow up was available. However, as described in Appendix 6 of the CS, a R-chemo+R cohort had to be constructed from patient level data for patients randomised to maintenance (PRIMA patients), that allowed estimates for PFS and PPS from the start of R-chemo induction therapy (as in the GALLIUM trial).

The transition probabilities used in the model are discussed in more detail below.

Probability of discontinuing from treatment

Time on treatment was directly estimated from the Kaplan-Meier (KM) curves for time-to-treatment-discontinuation (TTTD) obtained from the GALLIUM trial for both treatment arms. These are presented in Figure 5.2 and were used to estimate the proportion of patients on treatment in each cycle of the model. Parametric extrapolation was not needed since all patients in the GALLIUM trial had completed their treatment in both arms.

Figure 5.2: KM curves time to treatment discontinuation in the GALLIUM trial



Source: Figure 6 in clarification response.[8]

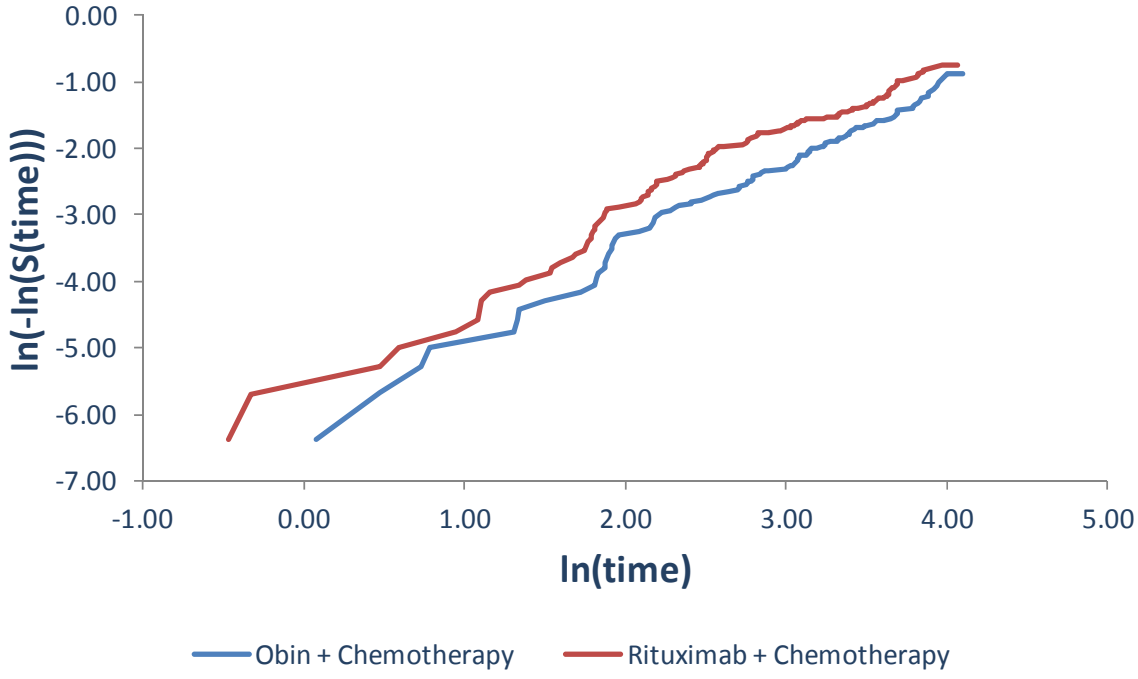
Progression free survival probability

The probability of remaining in PFS was estimated by fitting parametric probability distribution functions to the patient level PFS-INV data from the GALLIUM trial. This was done following the NICE Decision Support Unit (DSU) guidance.[40]

Proportional hazards assumption of PFS parametric functions

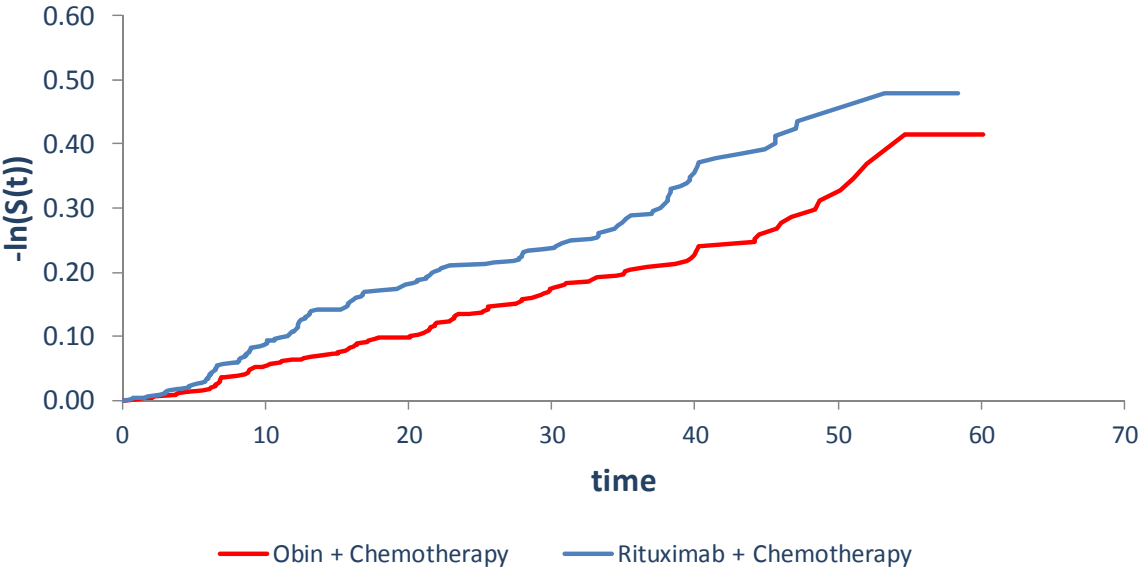
First, the proportional hazards assumption was checked and deemed valid after visual inspection of the log-cumulative hazards plot in Figure 5.3 and the cumulative hazard plot in Figure 5.4. For the extrapolation of PFS beyond the trial period, parametric functions were fitted simultaneously for both treatment arms data, with treatment as a covariate in the model.

Figure 5.3: Log-cumulative hazard plot for PFS in the GALLIUM trial (ITT FL population)



Source: electronic model in the clarification response.[41]

Figure 5.4: Cumulative hazard plot GALLIUM PFS INV – FL ITT



Source: electronic model in the clarification response.[41]

Goodness of fit of the PFS parametric functions

In order to choose the most suitable probability distribution function to model PFS, the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were calculated. The results are shown in Table 5.5. The lognormal, log-logistic or generalised gamma distributions presented the overall best fit according to AIC or BIC values.

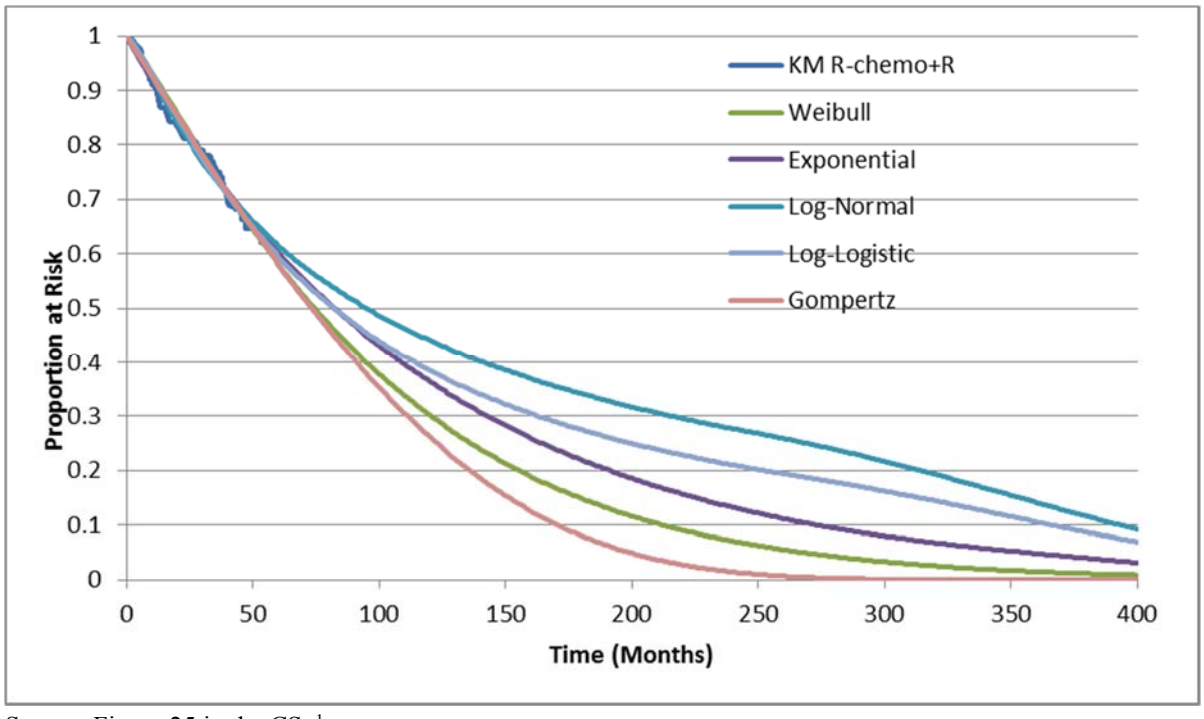
Table 5.5: Parametric functions, AIC and BIC goodness of fit for PFS

Distribution	AIC	Ranking	BIC	Ranking
Exponential	1785.9	5	1796.1	3
Weibull	1782.2	4	1797.5	5
Log-logistic	1779.9	3	1795.1	2
Lognormal	1774.5	1	1789.7	1
Generalised Gamma	1776.4	2	1796.8	4
Gompertz	1785.9	6	1801.2	6

Source: Table 63 in the CS. [1]
AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion

However, since the PFS data were immature in the GALLIUM trial, the company did not rule out the Exponential, Weibull or Gompertz distributions, also because these showed plausible fits to the observed GALLIUM data, as can be seen in Figure 5.5. However, all these survival functions differed in their long-term predictions of PFS. Thus, visual inspection and external validity of the tail of the PFS curve for the R-chemo+R arm was sought to justify the choice of the PFS parametric distribution. Note that external validation is only possible in the R-chemo+R arm as there are no long-term data available for the obin-chemo+obin arm.

Figure 5.5: PFS extrapolations, R-Chemo+R arm in the GALLIUM trial (FL ITT population – PFS-INV data)



Source: Figure 25 in the CS. [1]

To further select plausible parametric functions for PFS extrapolation, these were compared to available long-term data for the comparator R-chemo+R arm from the PRIMA trial and a publication from the US LymphoCare registry.[39, 42, 43] PRIMA is the main Phase III, randomised controlled trial of rituximab maintenance in patients with high tumour burden FL responding to rituximab plus chemotherapy induction. The follow up data were available for up to 9.75 years of an R-chemo+R cohort. Details of the PRIMA study and the analysis of PFS, PPS and OS for the R-chemo+R group are described in

Appendix 6 of the CS. The main limitation of the PRIMA study is that, at the time it was conducted, bendamustine was not available and therefore only data for patients receiving CHOP or CVP in induction was available for comparison. The US LymphoCare registry reports outcomes for US patients enrolled in LymphoCare with stage III/IV follicular lymphoma receiving R-CHOP, R-CVP or R with a fludarabine-based regimen (R-Flu) as frontline therapy.[43] Thus, similar to PRIMA, LymphoCare did not present long-term follow up data on bendamustine. The median follow-up was 7.4 years. The main limitation of the LymphoCare data is that not all patients potentially eligible for maintenance may have received maintenance as the registry enrolled prior to the wider use of maintenance after first-line induction.

Long-term PFS extrapolations of the different probability functions fitted to the GALLIUM R-chemo+R arm and the observed KM from PRIMA (R-chemo+R) and are shown in Figure 5.6 and PFS rates at different time points in Table 5.6. Based on these, the company concluded that, within the range of observed PFS behaviour, the exponential and log-logistic distributions seem to predict PFS rates in the observed range. The log-normal and generalised Gamma distributions seem to predict PFS at the high end and the Weibull distribution at the lower end. The Gompertz distribution seems to underestimate the observed PFS and was ruled out.

Figure 5.6: PFS extrapolations for R-chemo+R arm

Source: Figure 26 in the CS.[1]

Table 5.6: PFS rates at different time points for parametric functions (PFS-INV data)

	PFS at 6yrs (%)	PFS at 8yrs (%)	PFS at 10yrs (%)	PFS at 15yrs (%)
Exponential	54.6	44.6	36.4	22.0
Weibull	51.3	39.6	30.2	14.9
Log-logistic	54.1	45.2	38.5	27.5
Log-normal	57.1	49.8	44.1	34.2
Generalized Gamma	56.8	49.3	43.5	33.3
Gompertz	50.8	37.4	26.2	8.1

Source: Table 64 in the CS.[1]
PFS = progression free survival

Furthermore, based on the recommendations of an UK advisory board, a function representing the mid-range of plausible estimates was chosen. This constrained the choice to the exponential or log-logistic distributions only. For the base-case, the company selected the exponential distribution to model PFS. Nevertheless, alternative distributions were built in the model and investigated in sensitivity analyses. The base-case parameters for the exponential distribution are shown in Table 5.7.

Table 5.7: PFS base-case PFS fit parameters and covariance matrix – exponential distribution

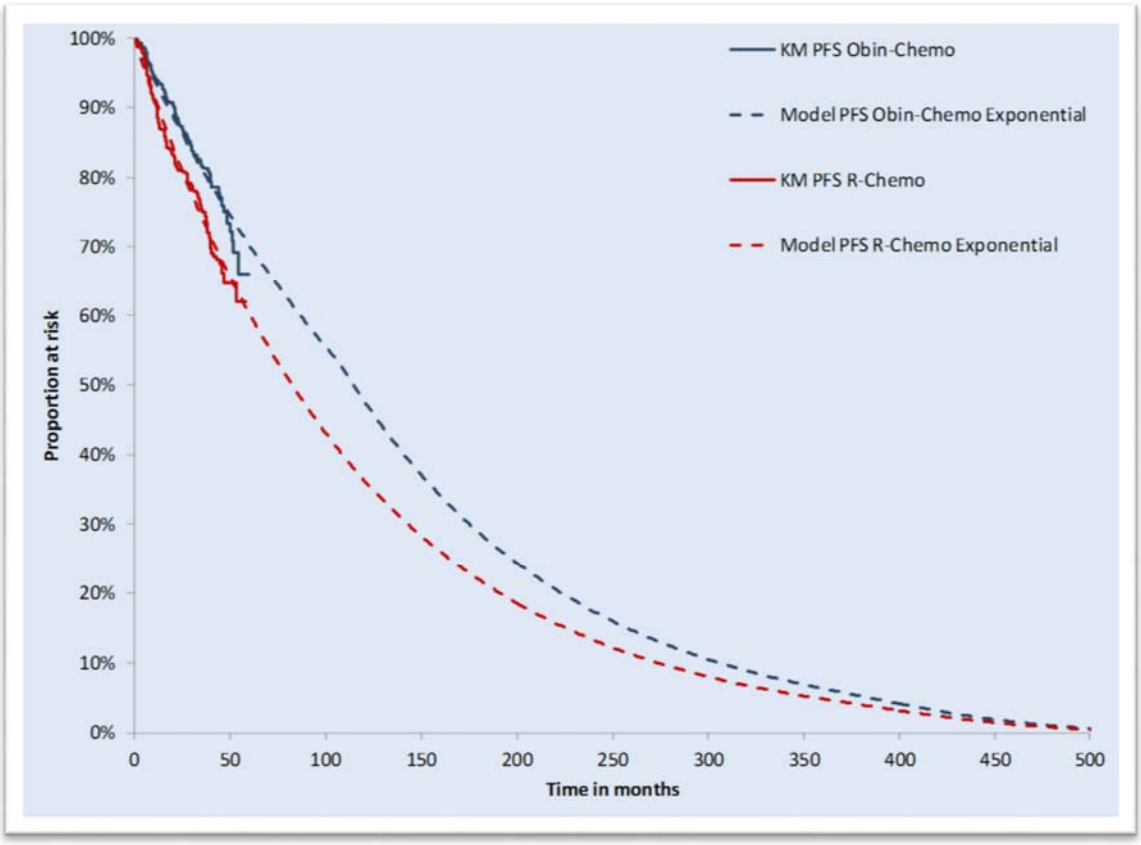
Fit		Covariance	
Parameter		**Intercept**	**Treatment**
Intercept	5.135	0.0083	-0.0083
Treatment (R-chemo)	-0.358	-0.0083	0.0145

Source: Table 65 in the CS.[1]

Long-term PFS on obin-chemo+obin

Long-term PFS on obin-chemo+obin was modelled in the base-case assuming proportional hazards (from GALLIUM) with treatment effect duration of 9.75 years. After this time point, a hazard of one (i.e. no treatment effect) was assumed. The assumption on treatment effect duration was based on the PRIMA study, where no indication of a finite duration of treatment effect on PFS was observed, and the proportional hazard assumption seemed to hold for the entire observation period (longest follow up 9.75 years). Furthermore, clinical advisors suggested that there is no evidence of a finite duration of treatment effect in treatments of FL and that it is plausible that this will be the case for obin-chemo+obin versus R-chemo+R. The PFS extrapolation assumed for the base-case is shown in Figure 5.7.

Figure 5.7: PFS base-case extrapolation



Source: electronic model in the clarification response.[41]

Probability of transitioning from PFS to death

The health economic model considers the UK age-specific all-cause mortality rates and the PFS death rate observed in the GALLIUM trial and uses the greater value of the two rates to determine the proportion of patients transitioning PFS to death. The monthly PFS mortality rate in the GALLIUM trial was calculated by pooling the number of deaths and the number of patient-months at risk in PFS between the arms. The pooled and per-treatment-arm figures are shown in Table 5.8.

Table 5.8: PFS death events and monthly death rates in the GALLIUM trial ITT FL (September 2016 cut-off date)

	N	Events	Patient months at risk	Monthly rate (95%CI)
Pooled	1202	38	39,519	0.096% (0.070%-0.132%)
Obin-chemo	601	23	20,389	0.113% (0.075%-0.170%)
R-chemo	601	15	19,130	0.078% (0.047%-0.130%)
Source: Table 6 in the clarification response CS. [8]				

Post progression survival probability

The model has two progressed disease states for early (progression occurs before two years) and late progression after first initial symptomatic treatment. Data from the GALLIUM trial was used to estimate the early PD mortality (there were no PPS events in late progression observed in the GALLIUM trial). Late PD mortality was estimated using data from the PRIMA trial. Monthly early post progression mortality rates were estimated from the individual patient level data from the

GALLIUM study. Due to the indolent nature of the disease, the data were immature and a relatively small number of events were available for analyses. For that reason, the data were analysed by pooling the treatment arms. For the PRIMA data, the R-chemo+R cohort was analysed as described in Appendix 6 of the CS. PPS KM curves for the early and late PD data sets from PRIMA are shown in Figure 5.8.

Figure 5.8: PPS in PRIMA for early PD (within two years) vs late PD (subsequent years)

Source: Figure 28 in the CS. [1]

Monthly mortality rates used in the base-case and sensitivity analyses in the model are shown in Table 5.9. The greater of the UK general population and the cohort mortality rates is applied to the transition probability from PD to death to account for the expected increase in long-term mortality due to age. Since data from PRIMA were not stratified by early and late progression, a pooled rate for the transition probability from PD to death was used for a scenario analyses.

Table 5.9: Monthly death rates in PD

	GALLIUM	PRIMA	PRIMA POOLED
Early progression (<2yrs)	1.61%	0.93%	0.77%
Late progression (>2yrs)	-	0.56%	0.77%
Source: Table 67 in the CS. [1]			

Transition rates (derived by fitting an exponential model to PPS curves) are shown in Table 5.10. No events were reported in the Late PD health state.

Table 5.10: PPS – Early PD (GALLIUM, FL ITT, September 2016 cut-off date)

	Number of Patients	Events	Monthly Rate
R-chemo+R	98	39	1.72%
Obin-chemo+obin	57	19	1.45%
Pooled	155	58	1.61%

Source: Table 7 in the clarification response.[8]

Overall survival probability

Overall survival (OS) is calculated in the model as the sum of the time spent in the PFS, early or late PD health states.

ERG comment: The ERG agrees with the company's approach of using GALLIUM data to estimate TTTD, PFS and PPS for early progressed disease, and PRIMA data to estimate PPS for late progressed disease. The analyses of the survival data were, in general, correctly performed and followed the guidance from the NICE Decision Support Unit (DSU).[40] The ERG's main concerns are explained in detail below.

Validity of the proportional hazards (PH) assumption: The company assessed the PH assumption visually (see Figure 5.3 and 5.4) and it was deemed valid. However, the ERG considers that assuming PHs based on this figure depends on how one interprets the plot and that a formal statistical test would have supported or rejected the choice of PHs. Such a test was provided in the clarification response (question B11). Given its result, the assumption of a PH model cannot be rejected with the current data. In any case, the company did not model the survival curves in the (standard) way that the ERG would have expected when the treatment effect is characterised with PHs, which is $S_{R-chemo}(t) = S_{Obin-chemo}(t)^{HR\ (R-chemo\ vs.\ Obin-chemo)}$. Since the company modelled PFS with regression equations including treatment effect as the only covariate and non-PH survival models (like the lognormal or log-logistic) were included in their analyses, the ERG considers that it was not necessary to test the PH assumption (because it is not explicitly used in the model).

Choice of PFS probability distribution: The ERG considers that the reason to choose between an exponential and a log-logistic distribution is unclear. While in the CS it is mentioned that the UK advisory board recommended using a function representing the mid-range of plausible estimates, it is unclear whether this was simply chosen by being in the middle or because it was validated based on clinical experience. In the clarification response (question B9),[8] the company explained that the clinical experts suggested that approximately 60-70% of patients would relapse within 10 years. Based on the predicted 10-year PFS probabilities shown in Table 5.6 above, the exponential and the log-logistic distribution seem appropriate choices for PFS since for both distributions the survival probability is between 30-40%. The ERG agrees with the company that in this case the exponential distribution is a more conservative choice and therefore preferred over the log-logistic. However, the ERG considers that the same reasoning is valid for the Weibull distribution. The 10-year PFS probability for the Weibull distribution is 30.2%. Thus, it is in the lower end of the values given by the experts. Furthermore, AIC and BIC values for the Weibull distribution in Table 5.5 are similar to those obtained for the exponential distribution. Therefore, the ERG considers that the exclusion of the Weibull distribution from the potential candidates to model PFS was not properly justified.

Treatment effect duration: The assumption of finite duration of treatment effect on PFS seems to be the main driver of the cost effectiveness results. In the absence of long-term data in the GALLIUM trial, this assumption was made based on data from the PRIMA trial. The CS states that "*in the PRIMA study there was no indication of a finite duration of treatment effect on PFS*" and thus "*the proportional hazard assumption for PFS seemed to hold for the entire observation period with longest follow up reaching of up to 9.75 years*".[1] However, it should be noted that the PRIMA trial compares rituximab maintenance after induction chemotherapy with observation (i.e. no maintenance), while in the GALLIUM trial maintenance with obinutuzumab versus rituximab are considered. In the clarification response (question B7),[8] the company states that "*as the mechanism of action of obinutuzumab as*

antiCD20 antibody is similar to that of rituximab, it is expected that the long-term effects of treatment observed with rituximab apply to obinutuzumab as well". While the ERG has no reasons to disagree with this statement, it should be pointed out that such long-term treatment effect would apply to obinutuzumab compared to observation (as in PRIMA). Whether the same long-term treatment effect applies to obinutuzumab when compared to rituximab is therefore speculative. The CS and the clarification response state that clinical advisors and the literature suggest that there is no evidence of a finite duration of treatment effect in treatments of FL and "*that it is plausible that this will be the case for Obin-chemo+Obin versus R-chemo+R*".[1, 8] However, it should be noted that there is no evidence of the opposite either (possibly due to the limited long-term follow-up data). According to the evidence in the company submission, the ERG could not propose an alternative estimate for the treatment effect duration that could have been considered robust. For that reason, the ERG explored the impact of this parameter in a threshold analysis. Based on its results, the ERG considers that assuming a treatment effect for five years (which also coincides with the longest follow up in the GALLIUM trial) could have been seen as a more conservative approach than the one presented in the company's base-case.

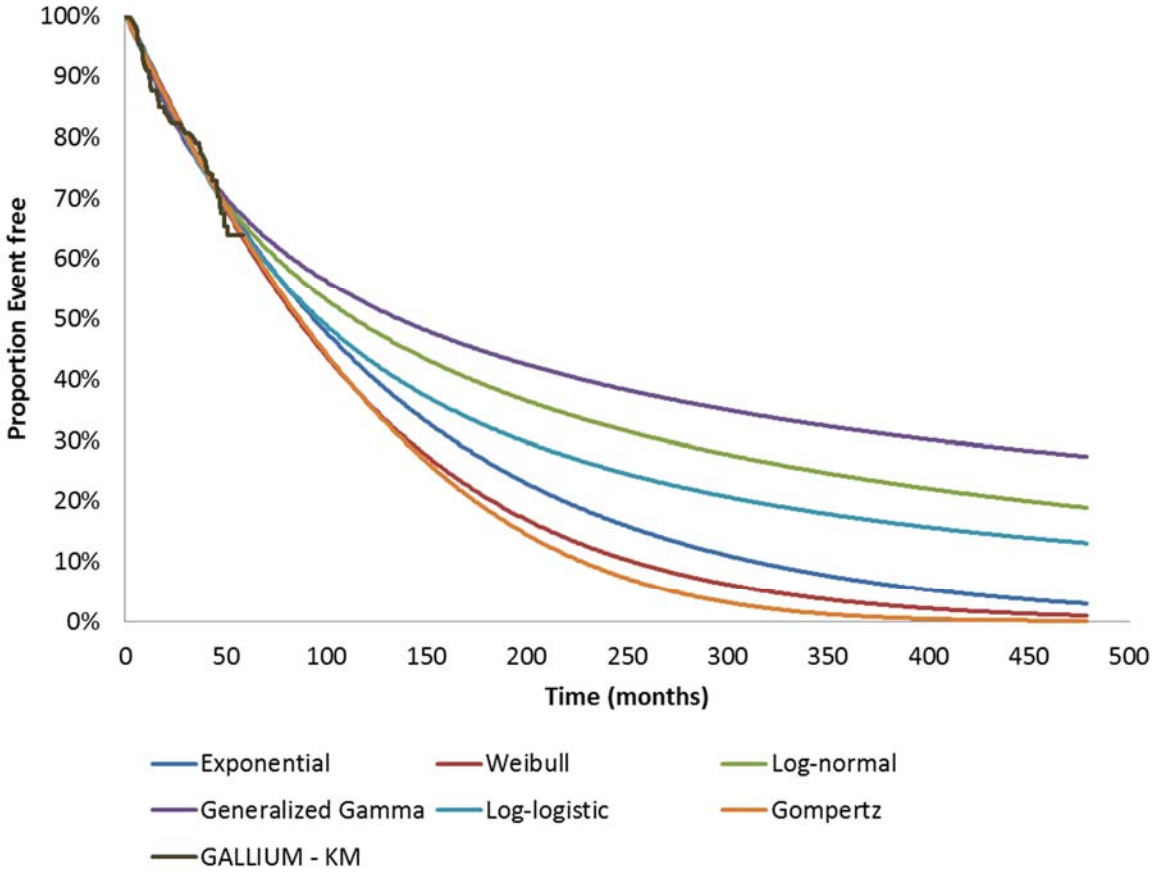
Transition probability from PFS to death: Page 147 of the CS states that the "*probability of death in PFS was derived from the observed mortality in PFS in the GALLIUM study. Since there were few events, number of deaths and the number of patient-months at risk in PFS were pooled between the arms*".[1] This implies that the probability of death in PFS was assumed equal for both treatment arms. However, this does not seem to be in line with the figures reported in Tables 28 and 29 of the CS, where the number of deaths observed in the obin and R arm were 21 (20.8% of the events) and 14 (9.7% of the events) and 24 (25.8% of the events) and 19 (15.2% of the events), respectively. Thus, it seems that the number of deaths during PFS was higher in the obinutuzumab arm. The ERG asked the company to present the number of events, patient-months at risk and monthly rates per treatment arm. These can be seen in Table 5.8 above (cut-off date September 2016) and it seems to confirm that indeed the number of deaths is higher in the obinutuzumab arm although the difference is not statistically significant. Nevertheless, when a treatment effect is sought; different values for the treatment arms should be used, regardless of statistical significance. Thus, in the ERG base-case analysis, different values per treatment arm will be considered.

Post progression survival: Page 149 of the CS states that the "*data were analysed by pooling the treatment arms and stratifying for early and late progression events*".[1] Pooling treatment arms can be considered correct if the number of events observed in both arms can be assumed to be the same. However, this was not reported in the CS. PPS data per treatment arm were provided in the clarification response (question B17).[8] The company analysed separately PPS for early and late progression. In late PD, there were no death events observed in either of the treatment arms. For early PD, a higher mortality rate was observed in the R-chemo+R arm although the difference was not statistically significant. In the ERG base-case analysis, different values per treatment arm will be used, since as mentioned above, the ERG considers that a treatment effect should be sought regardless of statistical significance.

Choice of PFS data: The investigator (INV) assessed PFS data (PFS-INV) were used in the base-case analysis. The ERG considers that independent review committee (IRC) assessed PFS (PFS-IRC) data should have been used for the company's base-case analysis because, as mentioned in the critique of Section 4.6 of this report, the GALLIUM trial was open-label. Therefore, results based on independent review are less prone to bias than investigator results. From a cost effectiveness analysis point of view, it would also represent a more conservative approach since the IRC analysis reported a higher hazard ratio for PFS than the one reported by the local investigator, meaning that using IRC data would imply less PFS benefit for obinutuzumab than the one using the local investigator data. The use of PFS-IRC data was investigated in a scenario analysis by the company. In this scenario, the company assumed the

same parametric distribution for PFS as in the company's base case: the exponential. However, the ERG considers that this is not correct. Since the PFS-IRC dataset is different from the PFS-INV dataset, the goodness of fit for the PFS-IRC data should have been reassessed. The ERG could not find in any of the documents submitted by the company AIC or BIC values for the different parametric functions when PFS-IRC data were chosen. Also, from the company's electronic model, it was not possible to plot the model's PFS-IRC extrapolated curves and compare them with the KM-IRC curves. However, the comparison for the R-chemo+R arm was done according to the KM-IRC curve provided by the company. This can be seen in Figure 5.9 below. Visual inspection of Figure 5.9 shows that all the parametric curves for the R-chemo+R arm present a similar fit to the GALLIUM data.

Figure 5.9: PFS extrapolations, R-Chemo+R arm in the GALLIUM trial (FL ITT population – PFS-IRC data)



Source: electronic model in the clarification response.[13]

Different PFS-IRC rates at different time points for the probability functions fitted to the GALLIUM R-chemo+R arm are presented in Table 5.11. Based on these, and in the validation criterion used by the company where approximately 60-70% of patients would relapse within 10 years, the ERG concluded that, within the range of observed PFS behaviour, the Weibull and Gompertz distributions seem to predict PFS rates in the observed range. With the information currently available, the ERG finds it difficult to make a choice between these two distributions. Since the ERG suggested that the company's base-case might have been based on the Weibull distribution for the local investigator assessment, it might be convenient to choose the Weibull distribution also in this case simply to facilitate a comparison between the two scenarios. In any case, other options should be explored in scenario analyses.

Table 5.11: PFS rates at different time points for parametric functions (PFS-IRC data)

	PFS at 6yrs (%)	PFS at 8yrs (%)	PFS at 10yrs (%)	PFS at 15yrs (%)
Exponential	58.8	49.2	41.2	26.5
Weibull	56.4	45.4	36.4	20.5
Log-logistic	58.6	50.1	43.5	32.3
Log-normal	61.1	54.2	48.7	38.9
Generalized Gamma	62.9	57.1	52.5	44.4
Gompertz	57.1	45.9	36.3	18.6

Source: electronic model in the clarification response.[13]
PFS = progression free survival

5.2.7 Adverse events

Safety results of the GALLIUM trial are discussed in detail in Section 4.2.2.8 of this report. Adverse events (AEs) are included in the cost effectiveness model in the form of costs (in the base-case analysis) and disutilities (in scenario anysis). The operationalisation of adverse events in the model are further described in Section 5.2.8 and Section 5.2.9 of this report. This refers to the updated version of the economic model submitted by the company on 23 June 2017 which included [redacted] relating to the data-cut from 10 September 2016.

ERG comment: In the company submission, a threshold of 2% was applied to the serious adverse events to create a shortlist of the most relevant/frequent ones from the GALLIUM trial. It was not clear to the ERG why an arbitrary 2% threshold was chosen, as the justification for this was not given in the company submission. Furthermore, this threshold was applied separately for grade 3, grade 4 and grade 5 adverse events, which led to a situation where grade 3 of an adverse event might be in the list (e.g. pnemonia), whereas grade 4/5 of the same type of adverse event might not. Due to data and time limitations, the ERG did not apply this threshold to the pooled grade 3/4/5 adverse events to create a new list of adverse events to be included in the model. Instead of that, for each adverse event type that was listed in the company submission, the cost and disutilities for all grades higher than or equal to 3 were incorporated into the cost and QALY calculations of the electronic model (e.g. pnemonia grade 3/4/5 were included in the ERG's analyses even though, based on the 2% threshold rule, only pnemonia grade 3 was included in the company's analyses).

5.2.8 Health-related quality of life

EQ-5D utility scores were collected in the GALLIUM trial and used in the cost effectiveness model. These utilities were collected at baseline, during treatment, after treatment, at the last assessment prior to progression, and at the first assessment after progression. Given the lack of long-term data beyond progression, utilities from GALLIUM were used for the PFS health state only (covering induction, maintenance and observation). Furthermore, the EQ-5D health index showed no statistically significant overall difference between the obin-chemo+obin and R-chemo+R arms during both treatment and follow-up periods. Therefore, the same utility values for PFS were assumed in both treatment arms. Utilities for the PD health states were sourced from the literature. Two studies were of potential interest.[36, 44]

Wild et al. collected data on 222 patients with FL in eight UK centres.[36] Utilities were elicited from patients using the EQ-5D questionnaire and clinical data collected allowed allocation of patients to five health states: active disease - newly diagnosed, active disease – relapsed, partial response to therapy, complete response to therapy/remission and disease free. Measurements were pooled to derive PFS and PD utilities.

Bec et al. report EQ-5D scores in a cross-sectional study of iNHL patients across Europe collected in an on-line questionnaire.[44] The study included utility values for PFS and PD but data were collected from only 18 UK patients. Therefore, the study by Wild et al. was considered the most relevant to inform the PD health states.

Health state utilities

Utility values collected in the GALLIUM trial by health state and treatment arm are presented in Table 5.12. Differences were observed between the two arms although these were not statistically significant.

Table 5.12: GALLIUM EQ-5D utility values by health state and treatment arm

	obin-chemo+obin		R-chemo+R		Difference	
State	**Estimate**	**Std. Err.**	**Estimate**	**Std. Err.**	**Estimate**	**P-value**
Induction - off tx	0.765	0.032	0.779	0.031	-0.015	0.72
Induction - on tx	0.823	0.015	0.824	0.015	-0.002	0.84
Maintenance & follow-up - off tx	0.826	0.015	0.810	0.015	0.017	0.13
Maintenance & follow-up - on tx	0.834	0.015	0.828	0.014	0.006	0.54
Early progression <= 2yrs	0.767	0.026	0.782	0.022	-0.015	0.62
Late progression > 2yrs	0.820	0.033	0.810	0.030	0.010	0.80
Source: Table 8 in clarification response.[8]						

Base-case utility values are summarised in Table 5.13. In scenario analyses, the company explored the use of EQ-5D utility scores at progression from GALLIUM.

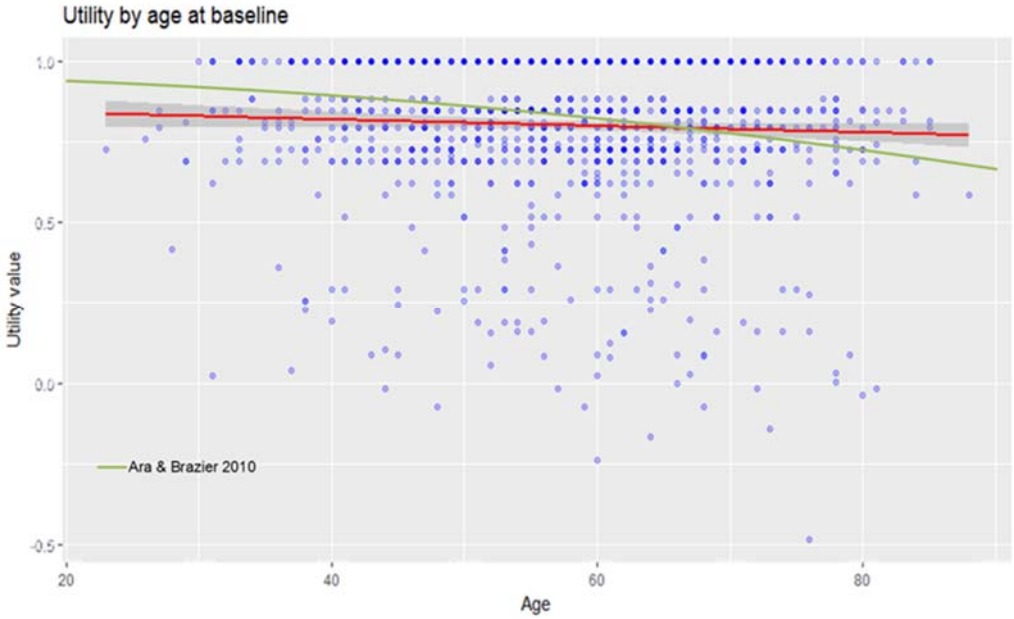
Table 5.13: Utility values for the base-case cost effectiveness analysis

Health state	Utility value: mean	Standard Error*	Reference in submission (section)	Justification
PFS (Induction - off tx)	0.772	0.027	Section 5.4.1 in the CS	GALLIUM trial based estimates
PFS (Induction - on tx)	0.823	0.007		
PFS (Maintenance & follow-up - off tx)	0.818	0.005		
PFS (Maintenance & follow-up - on tx)	0.831	0.006		
Early PD (including subsequent treatments)	0.62	0.06	Section 5.4.3 in the CS	Value from Wild et al. representative of later disease stage captured in the model progressed disease states.
Late PD (including subsequent treatments)	0.62	0.06	Section 5.4.3 in the CS	

Source: Table 73 in the CS.[1]
*See covariance matrix in Table 68 of the CS.[1]

In the base-case analysis, the company did not adjust the utility values for a decline in age. This assumption was justified on the basis of the EQ-5D values collected in the GALLIUM study at baseline, which according to the company, did not show an age dependent decline (cf. Figure 5.10). This assumption was tested in a scenario analysis, where the model utilities were adjusted for the age effects observed in the general UK population. [45]

Figure 5.10: GALLIUM baseline EQ-5D scores by age and UK general population.



Source: Figure 5 in clarification response.[8]

Adverse event disutilities

Disutilities for adverse events were not included in the base-case analysis. The company considered that this assumption was supported by the fact that the EQ-5D values were similar for patients on- and off-treatment in the GALLIUM trial and that the effects of AEs while on treatment may have been captured in the collected utilities. The company also considered that AEs were similar between the two treatment arms. Thus, including disutilities for AEs in the model would not result in a significant effect on the overall QALYs in each arm and the incremental difference between arms. Nevertheless, disutilities for AEs were considered in scenario analyses. Disutilities were sourced from the literature and were applied in the model for AEs of grade three and above that occurred in more than 2% of patients. The values used in the model can be seen in Table 5.14.

Table 5.14: Adverse event disutilities

Grade 3/4 adverse event	Disutility	SE	Source	Duration of adverse event (days)	Source
Neutropenia	-0.09	0.02	Nafees et al., 2008[46]	15.10	NICE TA 306[47]
Thrombocytopenia	-0.11	0.02*	Tolley et al., 2013[48]	23.20	NICE TA 306[47]
Anaemia	-0.12	0.02	Swinburn et al., 2010[49]	16.07	NICE TA 306[47]
Leukopenia	-0.12	0.02	Assumed to be same as Anaemia	16.07	Assumption
Pneumonia	-0.20	0.02	Beusterien et al., 2010[50]	14.00	NICE TA 306[47]

Source: Table 72 in the CS.[1]
SE = standard error.
*Assumed to be the average of all other adverse event disutility standard errors

ERG comment: The ERG's main concerns regarding the assumptions made for the utilities in the company's model are explained in detail below.

Literature results: Only two studies were deemed appropriate to source utilities by the company: Wild et al. (a conference abstract with two co-authors from Roche Products Ltd) and Bec et al. (a conference poster funded by Gilead).[36, 44] Table 70 in the CS[1] presents a summary of the studies that were deemed less applicable. In three studies it was unclear what the patient population was;[51-53] two other studies were single centre/small sample size studies, [54] [55] and a final study showed an unclear extrapolation from literature.[56] Thus, besides being a UK-based study, Wild et al. considered a relatively large sample of FL patients, the company argued.[36] However, the publication of Wild et al. to which the ERG had access to, was a poster abstract where no EQ-5D values were reported. A further search conducted by the ERG to retrieve the full publication was not successful. Therefore, the according utilities that were assumed in the CS could not be verified by the ERG. Nonetheless, the ERG found a study by Pettengell et al.[57] reporting HRQoL values of the FACT-Lym that seem to refer to the same study population as Wild et al. (Wild is listed as a co-author and the same number of included patients (N=222) are reported from eight sites in the UK). Using a mapping algorithm from Cheung et al.[58] to estimate UK utility values for the 'active disease relapse' group, which seems to correspond to the PD health state in the company's economic model, the ERG re-estimated the reported utilities by Wild et al. In addition, the ERG found EQ-5D utilities, based on the GADOLIN trial[59] that were reported in another, similar STA from the same company on obinutuzumab in combination with bendamustine for treating rituximab-refractory FL [ID841].[60] Comparing all these reported utilities (see Table 5.15), the ERG concluded that the newly estimated values were closer to the reported utilities of the GALLIUM and the GADOLIN trial and that the latter three were significantly different when compared to the utilities in Wild et al. or Bec et al. It should be noted that the mean utility values reported in the GALLIUM trial, GADOLIN and Wild et al. are similar for the PFS health state. The values of the PD states however vary extensively, ranging from 0.51 in Bec et al. (which reported lower values also for PFS) to 0.76 in GADOLIN.

Table 5.15: Literature-based mean and sources for utilities

Health state	Wild et al.[36]	Bec et al.[44]	GALLIUM[20]	GADOLIN[59]	Mapping FACT-Lym[57, 58]
PFS (on treatment)	0.81	0.71	0.82	0.82	NA
PFS (off treatment)			0.77	0.81	NA
PD	0.62	0.51	0.78 (early PD) 0.81(late PD)	0.76	0.73
PD = progressed disease; PFS = progression free survival					

Utilities for the PD health state: In spite of being unpublished, inconsistent with the results of the GALLIUM trial and unverifiable (by the ERG), the company relied on the utility values reported by Wild et al.[36] Based on the above-mentioned critique, the ERG judges the derivation and choice of EQ-5D utility values for the PD health state in the CS as non-transparent and non-replicable. However, given the uncertainties regarding the available the evidence, the ERG was not able to decide which of the estimates reported in Table 5.15 were the most reliable and representative for the patient population. For that reason, the ERG used the values from Wild et al. to model utilities in the PD health state in its preferred base-case and explored other alternative options in scenario analyses.

Choice of the GALLIUM data for the UK analysis: Page 152 of the CS[1] states that to "*inform the health state utilities in the economic model and to compare GALLIUM data to EQ-5D values to the literature, 5,007 observations from 1,097 patients were analysed with a mixed-effects model*". In light of this, it seems that the utilities used in the model were not constrained to UK values only and that instead all available data from GALLIUM were used regardless of geographical regions. However, in the CS it was not mentioned which tariff has been applied to the EQ-5D data in the GALLIUM trial, therefore it remains unclear whether a UK tariff has been applied to the GALLIUM utilities or not. Moreover, the commonly-used references for UK tariffs were not included in the CS either.[61, 62] The utilities reported in Table 5.12 and 5.13 could not be verified by the ERG. In the updated CSR document provided by the company,[63] the ERG could not find any UK-specific EQ-5D data. There are nevertheless several tables reporting EQ-5D values for Western Europe and it was observed that these seem to be lower than the overall ones. The reasons of the differences between utilities of Western European patients and others were not clear to the ERG.

Utilities for the PFS health state: The ERG agrees with the approach of the company of considering the utilities collected in the GALLIUM trial for the PFS health state only. It might be correct to use GALLIUM utilities for PFS and to assume that they are equal in both arms if the difference in means is not statistically significant. However, if a treatment effect is sought; different values for the arms should be used, regardless of statistical significance. Thus, the ERG preferred to use the utility values per treatment arm presented in Table 5.12 for its base-case but this option was not implemented in the company's model "*due to time constrains and the fact the utility differences in the GALLIUM trial (response to B19) were small and not statistically significant*". [64] Therefore, in the ERG's preferred base-case, the same pooled utilities per treatment arm were used for the PFS health state. While modelling different utilities per treatment arm is not expected to have a major impact on the ICER, it would result in an increased uncertainty around the ICER since more parameters would be included in the PSA.

Adverse event disutilities: The ERG agrees with the company that if the effects of adverse events while on treatment were captured in the collected utility values, then including additional disutilities would

not be necessary. However, the ERG considers that this is not completely clear. Throughout Section 4.2.2.8 of this report, the number of AEs reported per treatment arm show that these are more frequent in the obin-chemo+obin arm. The utility values per treatment arm presented in Table 5.12 show that, while on treatment, these utilities are nearly the same during the induction phase and higher for the obin-chemo+obin arm during maintenance. Thus, unless there is an additional HRQoL benefit of obinutuzumab compared to rituximab, which is not explicitly reported in the company submission, this seems unexpected (given the higher number of adverse events in the obinutuzumab arm).

Despite the ERG preference of using separate treatment arm specific utilities, the economic model only allows choosing pooled utilities for both treatment arms. Therefore, including AE disutilities can be regarded as an indirect way to reflect the difference in utilities between the two treatment arms. Under this approach, the ERG implicitly assumed that any difference in utilities between the treatment arms is due to the difference in adverse event rates. It should be noted that in the electronic model there are some minor inconsistencies between the list of adverse events that were considered to have a disutility implication and the list of adverse events considered to have a cost implication (e.g. febrile neutropenia is considered to have a cost implication but no disutility). While not incorporating the disutility of febrile neutropenia would result in an overestimation of the quality of life estimates, especially for the obinutuzumab arm, given the frequency of the adverse events, the impact on the ICER is expected to be minor. However, including these seemingly missing input parameters in the model would also increase the uncertainty around the ICER.

Different utilities for early and late PD: None of the potential sources considered for the PD health states utilities distinguished between the early and late PD health states. The ERG asked the company whether this was a reasonable assumption. In the response to the clarification letter (question B21), the company argued that, while this is plausible, they are not aware of any study reporting different utilities for patients progressing early or late.[8] The average utility values for patients progressing early in the GALLIUM trial were lower than those for patients progressing late. However, these values are also higher than the figures reported for PD in Wild et al.[36] The ERG agrees with the company that "*this may be due to the limited follow up in EQ-5D values beyond the point of progression in the GALLIUM trial leading to more censoring in patients progressing late*".[1] Different utilities for the early and late PD health states were explored by the ERG in scenario analyses.

Utility decrement with age: The ERG does not agree with the company's assumption of not adjusting the utility values for a decline in age. The interpretation of Figure 5.10 of the ERG report is subjective and while the company states that "*an age dependent decline is not observed*",[1] the ERG considers that in fact a slight and somewhat constant decline is shown. The ERG agrees with the company that "*it is not obvious that an age dependent decline observed in the general population should translate in the same way to a specific disease*", that the "*EQ-5D baseline values collected in the GALLIUM trial did not appear to be correlated with age*"[1] and that the decline for the general population seems inconsistent with the one observed in the GALLIUM trial, as shown in Figure 5.10. However, seeing the age distribution in the GALLIUM trial, it seems very unlikely that the trial was powered to detect differences in utilities for different age groups. Therefore, any assumption based on this does not seem to be valid. For that reason, the ERG considers that the decline in age for utilities should have been included in the base-case analysis, which would also result in a more conservative ICER.

5.2.9 Resources and costs

The company identified costs and health resources through a literature search. The full search strategy is reported in Appendix 5 of the CS[12]. Five studies fulfilled the in- and exclusion criteria of which four were already identified in the search for the cost effectiveness studies. [27, 28, 30, 31] The additional study

by Lewis et al. reported on a developed Markov model to evaluate the cost effectiveness of rituximab with CVP when compared to CVP alone. [65] Although no details on resource use and cost inputs were provided, lifetime healthcare costs per patients were reported. Most of the cost and resource inputs in the CS were not derived from any of the detected studies.

Health state costs

Health state related costs consisted of medication costs (induction and maintenance), supportive care costs, subsequent treatment costs in PD, and adverse event costs. Relevant medication costs included costs of obinutuzumab, bendamustine, CHOP, CVP, and rituximab. All drug acquisition costs were taken from the British National Formulary [66], or eMIT [67]. Different prices were considered separately for the induction and the maintenance phase.

Treatment with obinutuzumab

A list price of £3,312.00 ([redacted] with PAS) for a 1,000 mg vial of obinutuzumab was stated in the CS. In the GALLIUM trial, during induction, 1,000 mg obinutuzumab was administered by IV infusion together with bendamustine, CHOP, or CVP. In the obin-bendamustine regimen, obinutuzumab was administered in eight doses (Days 1, 8, and 15 of Cycle 1 and on Day 1 of Cycles 2-6). In the obin-CHOP and obin-CVP regimen, obinutuzumab was administered in 10 doses (Days 1, 8, and 15 of Cycle 1 and on Day 1 of Cycles 2-8). Patients responding to induction with obinutuzumab, received maintenance treatment with an IV infusion of 1,000 mg obinutuzumab every two months for up to two years or until disease progression.

Treatment with rituximab

The BNF list price for 100 mg and 500 mg of rituximab was stated with at £174.63 and £873.15 respectively.

[redacted]

During induction, in the GALLIUM trial, 375mg/m^2 rituximab were administered by IV infusion together with bendamustine, CHOP, or CVP. In the r-bendamustine regimen, rituximab was administered in six doses (Day 1 of Cycles 1-6). In the r-CHOP and r-CVP regimen rituximab was administered in eight doses (Day 1 of Cycles 1-8).

For maintenance therapy, patients responding to the rituximab induction therapy received one fixed dose of 375 mg/m^2 rituximab every two months for up to two years or until disease progression. Prices of rituximab are stated with £174.63 and £873.15 for a 100 mg and 500 mg vial, respectively. However, according to the CS, in England, about [redacted]% of eligible patients receive a sub-cutaneous formulation of rituximab in maintenance. The model assumes £[redacted] as a net price for one 1,400 mg vial of rituximab sub-cutaneous (SC), list price £1,344.65. During maintenance therapy, patients that responded to rituximab induction received 1,400 mg rituximab SC every two months for up to two years or until disease progression.

Chemotherapy regimens

In the model, bendamustine is used at 90 mg/m^2 on days 1 and 2 of each cycle with a price of £6.85 and £27.77 per vial of 25 mg and 100 mg, respectively.

For CVP and CHOP prices were based on eMIT data (version May 2016) [67]. The model uses the average actual dose of the GALLIUM trial which was consistent with a typical dosing, as stated in the CS. Assumed costs per vial, costs per cycle and dosing are summarised in Table 77 of the CS.

ERG comment: In the CS it is stated that the "*actual average doses reported in the GALLIUM trial were consistent with a typical dosing*".[1] It remains unclear whether this was only the case for cyclophosphamide, doxorubicin, vincristine and prednisolone. For bendamustine for instance, the actual typical dosing is stated with 120 mg/m^2 instead of 90 mg/m^2.[68] The ERG could not correct this potential error in the economic model as the cumulative dose in each cycle used in the model was hardcoded. In principle, a higher dose of bendamustine would increase the total drug acquisition costs. However, since the price of bendamustine is low compared to other drug prices used in the model, and since bendamustine is given at the same dose to a comparable number of patients in both treatment arms (345 in obin-chemo vs. 341 in R-chemo), the impact of this potential error on the ICER is expected to be minor.

Furthermore, the ERG noticed that in the (original and updated) company's model, for the obin+CVP regimen, both actual and planned dose for obinutuzumab were missing (see model sheet 'Dosing Calc' cells AN 16:146). Due to lack of data, the ERG could not incorporate the actual or planned dose into the cost calculations of the model, and assumed that exactly one vial is used at each cycle when obinutuzumab is administered for the obin+CVP regimen. Not using the actual/planned dose data for obinutuzumab for the obin+CVP regimen is favourable for the obinutuzumab arm, especially under "no vial sharing" assumption. Since the number of vials used at each cycle is assumed to be the minimum possible (i.e. just one vial), the total drug acquisition costs of obinutuzumab might have been underestimated. Incorporating the actual/planned dose for obinutuzumab for the obin+CVP regimen might increase the ICER.

Drug administration costs

All drug administration costs used in the model were based on NHS reference costs tariffs.[69] Pharmacy costs for preparation of the infusions and patient NHS transport costs were considered as well and based on Papaioannou et al. [28] and Curtis.[70]

The administration schedules per cycle and the applicable costs for R-chemo+R and obin-chemo+obin are depicted in Table 77 of the CS.[1]

ERG comment: The company provided a solid overview on the drug costs considered in the model. The ERG verified the references of all available sources. Although costs for the different treatment scenarios were provided in the CS, several things were unclear. For the administration costs shown in Table 79 of the CS the ERG asked for a full derivation of these costs per cycle. Based on the administration costs and the administration schedule described in the CS [1], the company provided a new table with a per cycle derivation of the administration costs.

Furthermore, it was unclear how costs would differ between the different cycles and according to the different length of the chemotherapeutic regimens. Therefore, the ERG asked the company to detail the costs per cycle. The company clarified this with Tables 9 and 10 of the clarification response[8] and presented drug and administration costs per cycle for the chemotherapy regimens in combination with obinutuzumab and rituximab respectively. Drug costs were based on the average actual administered dose and the acquisition costs as stated in the CS.[1]

In addition, in the CS it is noted that with *"Papaioannou et al. 30% of patients were assumed to require NHS transportation".*[1] The ERG noticed that, for the very same assumption, Papaioannou et al. are in

turn referencing a pervious NICE submission from the company.[71] The ERG judges this way of referencing as not transparent.

Supportive care costs

In the absence of UK data or guidelines, costs and frequencies of supportive care in the CS were based on ESMO guidelines [72] and the literature [28]. Frequency and intensity of haematologist visits for both induction (0-6 months) and follow-up (6-30 months) period were assumed to be different for patients in the progression free health state when compared to patients that progressed early or late. Likewise, follow-up visits were considered. Costs for these visits were based on Papaioannou et al.[28] For both induction and follow-up period one CT scan was assumed respectively. The average monthly costs for the PFS and PD health states are summarised in Table 81 of the CS.[1]

ERG comment: Frequencies of visits of supportive care are not always clearly stated and do not correspond to the frequencies reported in Table 81 of the CS, which correspond to those implemented in the model.

Subsequent treatment costs in PD

Average costs for subsequent treatment in the early and late PD states were included in the model. According to the CS, clinical advisors suggested that next line treatment choices post progression would be the same between the two arms and that treatment for early and later progressors would not significantly differ.[1] Since data on time to next anti-lymphoma treatment (NALT) from the GALLIUM trial were immature and heavily censored, subsequent treatment costs were based on literature values reported by Papaioannou et al.[28] Accordingly, total costs of £13,427 for subsequent treatments were assumed for early and late progression in both arms of the model. For the sensitivity analysis costs of £5,437.61 were assumed. These costs were based on NALT data from the GALLIUM trial (calculated in Appendix 7 of the CS.[12])

Adverse event cost

AEs of Grades 3, 4 or 5 were considered when they occurred in more than 2% of patients in either arm of the GALLIUM trial. For the model AEs were assumed to occur at a constant rate while on treatment. Unit costs assumed per event and grade considered are depicted in Table 82 of the CS, together with the respective references. Average monthly AE costs while on treatment were estimated to be £53.62 in the obin-chemo+obin arm and £45.85 R-chemo+R arm, respectively. Adverse event costs associated with subsequent treatment lines after progression were assumed to be included in the costs of subsequent treatments. For an overview of considered AEs and their respective unit costs see Table 5.16.

Table 5.16: Adverse event costs

Event (Grade)	Unit Cost	Reference
Anaemia (3)	£2,117	SA03G (NL)
Febrile Neutropenia (3)	£6226.29	NICE CG NHL, 2016
Dyspnea (3)	£0.00	Not costed
Infusion related reaction (3)	£600.65	SA31E (NS)
Infusion related reaction (4)	£600.65	SA31E (NS)
Neutropenia (3)	£867.00	LRiG estimate rev. TA162, TA175
Neutropenia (4)	£867.00	LRiG estimate rev. TA162, TA175
Pneumonia (3)	£4154.97	DZ11P (NL)
Leukopenia (3)	£3236.25	SA31E (NL)
Leukopenia (4)	£3236.25	SA31E (NL)
Thrombocytopenia (3)	£3236.25	SA31E (NL)
Thrombocytopenia (4)	£3236.25	SA31E (NL)
Source: Based on Table 82 in the CS[12] *NHS reference costs 2015-16; NL, non-elective long stay; NS, non-elective short stay		

ERG comment: The costs stated in Table 5.16 could be verified by the ERG for the most part. However, according to the National schedule of reference costs[73] the national average unit cost price for haemolytic anaemia with CC score 3 (SA03G) is £3,021 instead of £2,117. Also, thrombocytopenia is listed twice with grade 3 in Table 82 of CS with the same unit cost price. The ERG assumed this to be a textual mistake and instead considered the last entry as 'Thrombocytopenia (4)'. In the clarification response (question A5),[8] it is mentioned that [redacted]

5.2.10 Cost effectiveness results

Base-case incremental cost effectiveness analysis results

The results of the base-case scenario presented in this section are based on a revised version of the economic model submitted with the clarification responses. This version of the model includes the costs of [redacted] relating to the data-cut from 10 September 2016 that formed the basis of the economic analysis for the NICE submission on 10 May 2017. Additional changes to the base-case were the use of a longer time horizon of 50 years (as requested by the ERG) and the use of the latest UK general population life tables updated when implementing a longer time horizon.

The results of the base-case cost effectiveness analysis showed that obin-chemo+obin resulted in a total cost of [redacted] and 10.01 QALYs. The comparator, R-chemo+R, resulted in a total cost of [redacted] and 9.23 QALYs. Thus, obin-chemo+obin produced 0.78 additional QALYs at an incremental cost of [redacted] when compared to R-chemo+R, leading to an ICER of [redacted] The base-case cost effectiveness results are summarised in Table 5.17 below.

Table 5.17: Deterministic base-case results

Technologies	Total Costs (£)	Total LYG	Total QALYs	Inc Costs (£)	Inc LYG	Inc QALYs	ICER (£/QALY)
obin-chemo+obin	[redacted]	13.33	10.01	[redacted]	0.84	0.78	[redacted]
R-chemo+R	[redacted]	12.49	9.23	-	-	-	-

Source: Based on Table 1 Appendix A – clarification response[64]
Values in the table are discounted and half cycle corrected
ICER = incremental cost effectiveness ratio; LYG= life years gained; QALYs = quality adjusted life years.

The Markov trace by health states (PFS, PD and Death) is shown in Figure 5.11 below. Both arms show a similar trend. The percentage of the patients in PFS decreased in time, whereas the percentage of the patients in PD first increased and then declined. The percentage of dead patients increased over time but this was slower for the obin-chemo+obin arm. It can be observed that in the beginning the percentage of patients in PFS is 100%. As time increases, the difference in the proportion of PFS patients between the treatment arms increases too, since more patients stay in PFS in the obin-chemo+obin arm. However, as time exceeds 100 months (approximately) the two PFS curves begin to come closer. This is the result of the assumption of nine years duration of the treatment effect.

Figure 5.11: Markov trace for the base-case analysis

Source: electronic model in the clarification response.[13]

Disaggregated results of the base-case incremental cost effectiveness analysis

Disaggregated base-case QALYs and costs per health state are presented in Table 5.18 and Table 5.19, respectively. Based on these results, the company concluded that patients spend significantly longer average time in PFS, accounting for 79% of total absolute QALYs gained and less time in early PD (10% of absolute QALY gain) and late PD (11% of absolute QALY gain), respectively.

Table 5.18: Summary of discounted QALY gain by health state

	Obin-chemo +Obin	R-chemo +R	Difference	Absolute	% of absolute
Health state					
Progression free survival	7.20	6.13	1.07	1.07	79%
Progression < 2 yrs	0.28	0.42	-0.13	0.13	10%
Progression > 2 yrs	2.53	2.69	-0.15	0.15	11%
Total	10.01	9.23	0.78	1.36	100%

Source: electronic model in the clarification response[13]
Values in the table are discounted and half cycle corrected

Table 5.19: Summary of predicted resource use by category of cost

State	Cost (Obin-chemo)	Cost (R-chemo)	Cost difference	Absolute difference	% of absolute
PFS					
Obinutuzumab	[redacted]	0	[redacted]	[redacted]	[redacted]
Rituximab	0	[redacted]	[redacted]	[redacted]	[redacted]
Chemotherapy	371	365	5	5	[redacted]
Drug Administration	7,751	6,589	1,162	1,162	[redacted]
Adverse Events	1,274	1,037	237	237	[redacted]
Supportive Care	7,759	6,821	938	938	[redacted]
PFS Total	[redacted]	[redacted]	[redacted]		
Progressive disease					
Supportive care and subsequent treatment costs	10,3101	11,956	-1,646	1,646	[redacted]
Total PD & PFS	[redacted]	[redacted]	[redacted]	[redacted]	100%

Source: Based on Table 2 Appendix A – clarification response[64]
Values in the table are discounted and half cycle corrected.
PD = progressed disease; PFS = progression free survival

5.2.11 Sensitivity analyses

Probabilistic sensitivity analysis

A probabilistic sensitivity analysis (PSA) was carried out by the company to quantify the uncertainty associated to the deterministic cost effectiveness results. The model input parameters which were included in the PSA, with their corresponding probability distribution, are presented in Table 5.20. If possible, the uncertainty was characterised statistically, by a standard error or a covariance matrix. Otherwise, an error of 20% from the mean was assumed. Drug acquisition costs were kept fixed and therefore not included in the PSA.

Table 5.20: Parameters included in the probabilistic sensitivity analysis

Parameter	Uncertainty	Distribution
Parameters for PFS obin-chemo+obin/R-chemo+R arms	Covariance matrix Table 65 in the CS[1]	Multivariate normal
Probability of death in PFS	Standard Error Table 65 in the CS[1]	Log-normal
Probability of death in early PD	Covariance matrix in electronic model[13]	Multivariate normal
Probability of death in late PD	Covariance matrix in electronic model[13]	Multivariate normal
Utilities in PFS and PD states	Standard Error in electronic model[13]	Beta
Time on treatment	KM Greenwood CI in electronic model[13]	Log-normal
Administration costs	Standard Error in electronic model[13]	Log-normal
Pharmacy costs	20% of mean	Log-normal
Adverse event cost	20% of mean	Log-normal
Number of adverse events	Standard Error in electronic model[13]	Log-normal
Supportive care costs PFS & PD and subsequent treatments	20% of mean or Standard Error	Log-normal
Source: Table 89 in the CS.[1] PD = progressed disease; PFS = progression free survival		

The PSA results presented by the company were based on 1,000 model iterations. These were plotted on the cost effectiveness (CE) plane and shown in Figure 5.12. From this plot, a cost effectiveness acceptability curve (CEAC) was calculated and shown in Figure 5.13. A table presenting incremental costs and QALYs was not reported in CS, but it could be obtained from the electronic model (see Table 5.21). The probabilistic and deterministic results are comparable to the deterministic ones presented in Table 5.17. From the CEAC, it is estimated that the probability that obin-chemo+obin is cost effective compared to R-chemo+R is approximately [redacted] at a threshold of £30,000 per QALY gained.

Figure 5.12: Incremental cost and QALY PSA base-case results

Source: electronic model in the clarification response.[13]

Figure 5.13: Cost effectiveness acceptability curve

Source: Figure 2 in Appendix A - clarification response[13]

Table 5.21: Probabilistic base-case results

Technologies	Total Costs (£)	Total LYG	Total QALYs	Inc Costs (£)	Inc LYG	Inc QALYs	ICER (£/QALY)
obin-chemo+obin	[redacted]	13.26	9.98	[redacted]	0.82	0.76	[redacted]
R-chemo+R	[redacted]	12.44	9.21	-	-	-	-

Source: electronic model included in the clarification response[13]
Values in the table are discounted and half cycle corrected
ICER = incremental cost effectiveness ratio; LYG = life years gained; QALYs = quality adjusted life years.

Deterministic sensitivity analysis

A deterministic sensitivity analysis (DSA) was carried out by the company. Continuous parameters were varied using the 10 and 90%-percentile values obtained from the PSA. In addition, categorical variables were also changed: such as parametric functions for PFS; PPS source; different settings for the time-on treatment; vial sharing and administration costs. The discount rates for costs and outcomes were varied according to standard methods and the time horizon altered. The parameters included in the DSA, with their corresponding lower and upper limits and the resulting ICERs are presented in Table 5.22 and depicted graphically in a tornado diagram in Figure 5.14.

Table 5.22: Deterministic sensitivity analysis for base-case

Parameter modified	Base-case value	High value	Low value	ICER high	ICER low
Utilities					
Utility in PFS - Induction - On tx	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Utility in PFS - Induction - off tx	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Utility in PFS - Maintenance - off tx	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Utility in PFS - Maintenance - off tx	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Utility in PD - Early progression ≤ 2yrs	0.618	0.693	0.547	[redacted]	[redacted]
Utility in PD - Late progression > 2yrs	0.618	0.693	0.547	[redacted]	[redacted]
Utility source PFS	GALLIUM		Wild		[redacted]
Utility source PD	Wild		GALLIUM		[redacted]
Utility age adjusted	No		Yes		[redacted]
AE Utility included	No		Yes		[redacted]
Costs					
1st administration Obin-chemo	430	535	347	[redacted]	[redacted]
1st administration R-chemo	430	532	356	[redacted]	[redacted]
Administration Obin-chemo (subsequent)	384	423	348	[redacted]	[redacted]
Administration R-chemo (subsequent)	384	421	350	[redacted]	[redacted]

Parameter modified	Base-case value	High value	Low value	ICER high	ICER low
Administration maintenance Obin	360	454	287	[redacted]	[redacted]
Administration maintenance R	303	394	238	[redacted]	[redacted]
Supportive care PFS induction	253	292	223	[redacted]	[redacted]
Supportive care PFS maintenance	83	95	72	[redacted]	[redacted]
Supportive care PFS follow up	58	67	50	[redacted]	[redacted]
AEs - obin-chemo+obin	54	58	51	[redacted]	[redacted]
AEs - R-chemo+R	46	50	43	[redacted]	[redacted]
Supportive care early PD	231	272	200	[redacted]	[redacted]
Supportive care late PD	58	67	50	[redacted]	[redacted]
Subsequent treatment early PD	13,427	17,038	10,406	[redacted]	[redacted]
Subsequent treatment late PD	13,427	17,065	10,445	[redacted]	[redacted]
Subsequent treatment early/late PD	13,427		5,437.61		[redacted]
Vial sharing	Yes		No		[redacted]
Time on treatment	Actual treatment duration		According to label		[redacted]
Rituximab sub-cutaneous use	[redacted]%	80%	40%	[redacted]	[redacted]
Outcomes					
PFS Parametric distribution function	Exponential		Weibull		[redacted]
PFS Parametric distribution function	Exponential		Log-normal		[redacted]
PFS Parametric distribution function	Exponential		Generalised Gamma		[redacted]
PFS Parametric distribution function	Exponential		Log-logistic		[redacted]
PFS Parametric distribution function	Exponential		Gompertz		[redacted]
PFS data set	Investigator		IRC		[redacted]
PFS treatment effect	9 years	No finite duration	5 years	[redacted]	[redacted]
PPS early PD	GALLIUM	PRIMA			[redacted]
PPS early & late PD pooled	Early/late	PRIMA Pooled	GALLIUM Pooled		[redacted]
Discount rate cost & effect	3.50%		1.5%		[redacted]
Time horizon (years)	50		40		[redacted]

Source: Table 3 in Appendix A - clarification response[64]

* It was not correctly reported in the clarification response. The values are based on the updated version of the electronic model.[13]

AE = adverse event; ICER = incremental cost effectiveness ratio; PD = progressed disease; PFS = progression free survival; PPS = post progression survival; tx = treatment.

Figure 5.14: Tornado diagram for base-case

Source: Figure 3 in Appendix A - clarification response.[64]

From the DSA results, it was observed that the ICERs remained below [redacted] and were close to the base-case value in most cases. The most influential parameter was, as expected, the duration of the treatment effect, whose variation resulted in a wide range of possible values of the ICER. The ICER was also sensitive to the following parameters:

Clinical inputs

The ICER was sensitive to the choice of parametric distributions for PFS extrapolation. Assuming a log-normal parametric distribution for PFS resulted in a lower ICER of [redacted]/QALY whereas assuming a Weibull distribution increased the ICER to [redacted]/QALY. Assuming the IRC assessed PFS resulted in an ICER of [redacted]/QALY. Additionally, shortening the duration of the treatment effect to five years (longest follow up in the GALLIUM trial) resulted in an ICER of [redacted]/QALY. With no finite duration of the treatment effect on progression the ICER was [redacted]/QALY.

Utilities

The ICER was sensitive to the assumptions on the utility values used in the PD states: using the values from GALLIUM increased the ICER to [redacted]/QALY. Additionally, adjusting the utilities for a decline in age (in line with the general UK population) increased the ICER to [redacted]/QALY.

Cost

The ICER was sensitive to the drug acquisition costs. Using time on treatment as per protocol (i.e. assuming all patients in PFS would receive treatment per protocol while in PFS rather than as observed in the GALLIUM trial) increased the ICER to [redacted] due to the higher technology costs. However, this scenario would require full adherence to protocol in clinical practice and disregard treatment discontinuation due to tolerability or other reasons. In addition, the increase in ICER relies on the assumption that the additional treatment received per protocol would have resulted in no additional clinical benefit compared to the observed benefits in the GALLIUM trial.

Discounting

Due to the indolent nature of the disease, a significant amount of health benefits accrue over a longer time period. The ICER was therefore sensitive to the discount rate and using an alternative value of 1.5% (for costs and health effects) decreased the ICER significantly to [redacted]/QALY.

ERG comments: The ERG noticed some inconsistencies in the calculation of upper and lower bounds of the administration cost items (e.g. for some of the cost components, the upper bound of the cost component was lower than the average value). The ERG corrected these errors in its base-case.

Scenario analyses

The company conducted several scenario analyses to explore the impact on the cost effectiveness results of several of the structural uncertainties which are present in the economic evaluation. The company considered two scenarios in the CS (scenario analysis 1 and 2). Additionally, the company ran three scenarios as requested by the ERG in the clarification letter. These were reported in Appendix A from the clarification response.[64] The results of the cost effectiveness analyses are based on the updated version of the electronic model received with the clarification response.

Scenario analysis 1 - Alternative PFS and PPS assumptions

The main purpose of this scenario was to compare the company de novo model with the latest economic analysis of rituximab in combination with chemotherapy in the first-line treatment of FL conducted by Papaioannou or Dewilde. [28, 30] The main differences between these and the company's model can be found in the assumptions made for modelling PFS and PPS. A log-normal PFS extrapolation with no limit on the duration of treatment effect was used in Papaioannou et al.[28] Furthermore, PPS was not explicitly dependent on time to progression after first-line treatment (i.e. there is no distinction between early and late PD). Thus, in this scenario, the company considered a log-normal PFS extrapolation with no limit on the duration of treatment effect and assumed that the probability of death in early and late PD was the same (and equal to the pooled early and late post progression mortality observed in the PRIMA trial). Additionally, this scenario considered age-dependent reduction in utilities. Although under these assumptions an increased gain in life years in PFS (3.75 years median, 2.63 mean undiscounted) compared to the base-case (2.75 years median, 1.92 mean undiscounted) was achieved, the resulting overall life years gained (1.42 mean undiscounted) were similar to the base-case (1.45 mean undiscounted) and the resulting ICER, shown in Table 5.23, was also comparable to the base-case (Table 5.17).

Table 5.23: Scenario analysis – alternative PFS and PPS assumptions

Technologies	Total Costs (£)	Total LYG	Total QALYs	Inc Costs (£)	Inc LYG	Inc QALYs	ICER (£)
obin-chemo+obin	[redacted]	13.77	10.07	[redacted]	0.79	0.74	[redacted]
R-chemo+R	[redacted]	12.98	9.33	-	-	-	-

Source: Table 4 in Appendix A - clarification response[64]; Values in the table are discounted and half cycle corrected
ICER = incremental cost effectiveness ratio; LYG = life years gained; QALYs = quality adjusted life years.

Scenario analysis 2 - Assumptions of equal QALYs and costs post progression

In this scenario, the company assumed no difference in costs and QALYs gained post progression. This scenario was previously proposed by the ERG for TA251 in chronic myeloid leukaemia.[74] The resulting ICER, shown in Table 5.24, was lower than the ICER in the base-case scenario (Table 5.17).

Table 5.24: Scenario analysis – assumption on equal post progression QALY and cost

Technologies	PFS Costs (£)	Total LYG in PFS	Tot QALYs in PFS	Inc Costs (£)	Inc LYG	Inc QALYs	ICER (£)
obin-chemo+obin	[redacted]	8.78	7.20	[redacted]	1.31	1.07	[redacted]
R-chemo+R	[redacted]	7.47	6.13	-	-	-	-

Source: Table 4 in Appendix A - clarification response[64]; Values in the table are discounted and half cycle corrected
ICER = incremental cost effectiveness ratio; LYG = life years gained; QALYs = quality adjusted life years.

Scenario analysis 3 - Older starting age at baseline (clarification response questions B13 and B15)

The age at baseline of FL patients receiving chemotherapy in the HMRN database cohort was reported as a median of 63.7 which is 4.7 years older than the median age of patients in the GALLIUM trial (59.0 years). The impact of age in the cost effectiveness results was investigated by increasing the mean age at baseline in the model from 57.9 years to 62.6 years (i.e. by the difference in median as the HMRN report did not report mean). The results are shown in Table 5.25. The increase in the ICER can be attributed to the fact that an older cohort would gain less QALYs due to the reduced life expectancy.

Table 5.25: Scenario analysis – assumption on older first line FL treatment starting age

Technologies	Total Costs (£)	Total LYG	Tot QALYs	Inc Costs (£)	Inc LYG	Inc QALYs	ICER (£)
obin-chemo+obin	[redacted]	12.73	9.63	[redacted]	0.74	0.71	[redacted]
R-chemo+R	[redacted]	11.99	8.92	-	-	-	-

Source: Table 6 in Appendix A - clarification response.[64]
Values in the table are discounted and half cycle corrected
ICER = incremental cost effectiveness ratio; LYG = life years gained; QALYs = quality adjusted life years.

Scenario analysis 4 - Different chemotherapy distribution (clarification response questions B14 and B15)

Cost for chemotherapy, administration and adverse events were re-weighted according the distributions presented in Table 14 in the CS.[1] Based on all patients receiving bendamustine, CHOP or CVP (100%) the weights were: 37.18% for bendamustine, 16.67% CHOP and 46.15% CVP, respectively. This resulted in a smaller cost difference between the R-chemo+R and the obin-chemo+obin arm in drug acquisition, administration and AE costs, decreasing thus the ICER. A summary of the cost effectiveness results can be seen in Table 5.26.

Table 5.26: Scenario analysis – assumption different chemotherapy mix (cost only)

Technologies	Total Costs (£)	Total LYG	Tot QALYs	Inc Costs (£)	Inc LYG	Inc QALYs	ICER (£)
obin-chemo+obin	[redacted]	13.33	10.01	[redacted]	0.84	0.78	[redacted]
R-chemo+R	[redacted]	12.49	9.23	-	-	-	-

Source: Table 7 in Appendix A - clarification response. [64]
Values in the table are discounted and half cycle corrected
ICER = incremental cost effectiveness ratio; LYG = life years gained; QALYs = quality adjusted life years.

Scenario analysis 5 - Different PFS mortality and PPS by treatment arm (clarification response questions B12 and B17[8])

A separate mortality in PFS and PPS (early PD only) between the obin-chemo+obin and the R-chemo+R arm were included in the model. The point estimate for the mortality in PFS in the obin-chemo+obin arm was higher than in the R-chemo+R arm, but this was reversed in PPS, with a lower mortality estimate in the in the obin-chemo+obin arm. Note that the ERG found a programming error in the implementation of this scenario, and realised that the CS model still used the pooled death probability for obin-chemo+obin arm. This error was corrected by the ERG (as explained further in Section 5.3.1 in detail). A summary of the corrected cost effectiveness results for this scenario can be seen in Table 5.27. This scenario resulted in an ICER higher than the base-case ICER (obtained using pooled mortality rates in PFS and PPS) which shows the impact of an increased overall mortality in the obin-chemo+obin arm.

Table 5.27: Scenario analysis – PFS and PPS (Early PD) by treatment arm*

Technologies	Total Costs (£)	Total LYG	Tot QALYs	Inc Costs (£)	Inc LYG	Inc QALYs	ICER (£)
obin-chemo+obin	[redacted]	13.27	9.98	[redacted]	0.73	0.72	[redacted]
R-chemo+R	[redacted]	12.54	9.26				

Source: Table 8 in Appendix A - clarification response.[64]
Values in the table are discounted and half cycle corrected
*programming error corrected
ICER = incremental cost effectiveness ratio; LYG = life years gained; QALYs = quality adjusted life years.

ERG comment: The cost effectiveness analyses were correctly performed and well-presented in general. The ERG considers that assumptions regarding the choice of parametric functions for PFS or the PPS source, for example, should have not been included in the DSA but presented in the scenario analyses since these assumptions are concerned with structural (but not with parameter) uncertainty. While the ERG acknowledges that this is something minor, the ERG prefers to make this distinction to avoid for instance presenting a tornado diagram where in some cases there is no upper or lower limit. This will be applied to the results presented in Section 5.3.

Several structural uncertainties were tested by the company either in the DSA (as categorical variables) or as scenario analyses. With the exception of the first scenario analysis, all the uncertainties were explored individually and therefore a combined effect of multiple, and possibly conservative assumptions on the ICER, is missing. This will be explored by the ERG in Section 5.3.

Throughout this report, the ERG has drawn attention to a number of parameters for which it is believed that there is uncertainty and therefore, they should have been included in the probabilistic analyses. These parameters include the duration of the treatment effect, the Body Surface Area used to calculate drug dosages, the proportion of patients treated with bendamustine, CHOP or CVP, different utilities per treatment arm and the costs and disutilities for all adverse events of grade 3 or higher. This was explored by the ERG for the duration of the treatment effect and the costs and disutilities for pneumonia (grade >3). However, the other aforementioned parameters could not be included in the PSA due to lack of data and time constraints. While including all these parameters in the probabilistic analyses of the model is expected to have a minor/moderate impact on the ICER, the current probabilistic results are likely to underestimate the uncertainty around the model results. The ERG considers it difficult to estimate to what extent the uncertainty is underestimated. In the scenario where the duration of the treatment effect was included in the PSA, it was observed that the probability that obin-chemo+obin is cost effective was approximately [redacted] at the £30,000 threshold ICER. Thus, an absolute reduction of [redacted] and [redacted] compared to the ERG and company base-case probability at the same threshold, respectively. Including more parameters in the PSA might decrease the cost effectiveness probability even more.

5.2.12 Model validation and face validity check

In Section 5.10 of the CS, it is mentioned that the model concept with key clinical inputs, assumptions and clinical outputs was presented to a clinical advisory board of 9 UK clinicians to ensure face validity of assumptions and main clinical results, and that the electronic version of the model was checked by an external agency to guarantee the technical validity of the model. [1]

Face validity of the model's results can be assessed from the PFS and OS predicted by the model. These are shown in Figure 5.15 and Figure 5.16, respectively, and in Tables 5.28 and 5.29 below. The predicted PFS was in the range observed in long-term follow-up data for the R-chemo+R arm from the PRIMA trial, as discussed in Section 5.2.6 of this report. In addition, the median PFS in the R-chemo+R arm was in line with expectations of the clinical experts consulted by the company (Table 5.29).

The model OS estimates for 12, 24, 36 and 48 months are shown in the Table 5.28 in comparison with the KM estimates obtained from GALLIUM. The model seems to overestimate OS initially and then underestimate OS in both arms. In particular, this over- and underestimation is higher for the R-chemo+R arm. However, with the exception of the OS for the R-chemo+R arm at 12 months, the model estimates are within the KM 95% confidence intervals for both arms. Furthermore, the model predicts 20% fewer OS events in the obin-chemo+obin arm at 48 months. This represents a HR of 0.80 which is in line with the HR observed in the GALLIUM trial of 0.82 (95% CI: 0.54, 1.22). Thus, the model seems to predict a plausible difference in OS between the arms although slightly favourable to the obin-chemo+obin arm. In any case, due to the indolent nature of the disease, OS data in the GALLIUM trial was deemed immature. Therefore, it was not possible to validate the long-term OS predictions of the model against GALLIUM data. The clinical experts consulted by the company considered that the median predicted OS of 16.5 years for the R-chemo+R arm exceeded their expectations of about 14 years. However, the experts acknowledged that their current experience with long-term survivors was based on a cohort of patients who started treatment more than 10 years ago and that with the current standard of care it might be plausible to obtain a higher OS.

Table 5.28: Model OS prediction versus KM estimates

	obin-chemo+obin		R-Chemo+R	
Months	Model	KM (95% CI)	Model	KM (95% CI)
12	98.4%	97.8% (96.6%-99.0%)	98.2%	96.4% (94.9%-97.9%)
24	96.0%	95.5% (93.9%-97.2%)	95.1%	93.5% (91.5%-95.5%)
36	93.3%	93.9% (92.0% -95.9%)	91.7%	92.2% (90.0% -94.4%)
48	90.6%	91.5% (88.9%-94.2%)	88.3%	90.6% (88.1%-93.2%)

Source: Table 12 in clarification response (question B30).[8]
KM = Kaplan-Meier

Table 5.29: Base-case model PFS and OS outcomes (FL ITT, undiscounted)

	obin-chemo+obin	R-chemo+R	Difference
Mean LYs in PFS	11.60	9.68	1.92
Median PFS	9.58	6.83	2.75
Total Mean LYs (OS)	19.42	17.97	1.45
Median OS	18.67	16.50	2.17

Source: Table 86 in the CS.[1]
LYs = life years; OS = overall survival; PFS = progression free survival

Figure 5.15: Model base-case PFS and OS (FL ITT population)

Source: electronic model in the clarification response.[13]

Figure 5.16: Model OS (FL ITT population)

Source: electronic model in the clarification response.[13]

The company also compared the model predictions for the R-chemo+R arm to those from the models developed by Papaioannou et al. and Dewilde et al.[28, 30] These models predicted mean times in PFS for R-chemo+R from 5.2 years (R-CVP, Dewilde) to 8.5 years (R-benda+R, Dewilde) and mean OS from 11.7 years (R-CVP+R; Dewilde) to 13.1 years (R-benda+R, Dewilde). The company concluded that, whereas the PFS predictions are lower for R-chemo+R in these models, the current analysis predicts higher mean OS values in the R-chemo+R arm, and, therefore a longer time in post progression. Furthermore, the company argued that to obtain a mean OS value of approximately 14 years, higher post progression mortality should be assumed; but that this would contradict the observed mortality difference between early and late progression. The company considers that a more likely explanation is that previous models underestimated post progression survival, due to the data used to model these outcomes. The study by Papaioannou et al.[28] used data from van Oers et al.[75] but, in this cohort, approximately 50% of patients had less than two years from initial diagnosis and approximately 20% had two prior treatments. The company considered it likely that these patients had therefore progressed early and that post progression survival was underestimated.

ERG comment: The CS does not provide details about the design and results of the different validation exercises conducted by the company.

The ERG conducted some of the steps of an in-house technical verification checklist (TECH-VER checklist) to verify whether the model was implemented correctly and whether the report (description of the electronic model, inputs as well as results) and the electronic model (inputs, calculations and results) were consistent or not. The protocol steps and cell by cell checking of the calculations in the model helped the ERG in identifying a number of programming errors as well as reporting inconsistencies, which are corrected in the ERG exploratory analyses. The corrected errors/inconsistencies are explained in Section 5.3.1.

PFS validation of R-chemo+R arm with data form PRIMA seems to be fine. Comparing the median PFS values in both arms, a HR of approximately 0.7 could be assumed which is in line with the HR's reported in the CS.

In the absence of long-term OS data, the company's validation efforts focused on validation at face value with clinical experts. While it is mentioned that the predicted median OS for the R-chemo+R might be considered high by the experts, nothing is said about the obin-chemo+obin arm. If the clinical experts' estimate of 14 years for the median OS is considered valid, applying the model HR for OS of 0.80 would result in 16.8 years for the median OS in the obin-chemo+obin arm.

Visual inspection of Figure 5.15 and 5.16, makes it clear the difficulty of extrapolating OS with the current data from GALLIUM. While the extrapolated curves do not seem to match the observed data, this is probably due to the lack of OS events observed in the GALLIUM trial. For PFS however, the extrapolated curves seem to match quite well the KM curves from GALLIUM, especially for the R-chemo+R arm. In the obin-chemo+obin, the parametric and the KM curves show a good fit up to 50 months (approximately) but then, the parametric curve deviates from the KM curve overestimating the PFS probability. This seems to be caused by two main reasons: the duration of the treatment effect and the choice of investigator assessed PFS data. This will be explored in Section 5.3.

The comparison of the company's model predictions for the R-chemo+R arm to those from the previously published models by Papaioannou et al.[28] and by Dewilde et al.[30] revealed that the company model predicts higher mean PFS and OS estimates for the R-chemo+R arm compared to the reported results in these previously published models. While the company attempted to explain the potential causes of these differences, without the economic models and detailed description of each model, the ERG finds it impossible to pinpoint the exact reasons of the differences between the OS and PFS estimates of the studies.

5.3 Exploratory and sensitivity analyses undertaken by the ERG

After all the considerations discussed in Section 5.2, the ERG decided to define a new base-case scenario. This base-case scenario included multiple adjustments to the base-case analysis presented in the clarification response and summarised in Section 5.2. The adjustments made by the ERG for its base-case scenario were subdivided into the following three categories (according to Kaltenthaler et al.)[76]:

- Fixing errors (correcting the model where the company's electronic model was unequivocally wrong)
- Fixing violations (correcting the model where the ERG considered that the NICE reference case, scope or best practice has not been adhered to)
- Matters of judgement (amending the model where the ERG considers that reasonable alternative assumptions are preferred)

After the ERG base-case analysis, additional scenario analyses were performed in order to explore the impact of alternative assumptions on the cost effectiveness analyses results.

5.3.1. Explanation of the ERG adjustments

Fixing errors

1. Fixing errors consisted of:
    a. Correcting the PFS mortality rate for the obin-chemo+obin arm when treatment specific probabilities were used: In the updated version of the company's model, when treatment

specific death probabilities in PFS were chosen, the pooled monthly death probability of 0.096% was used for the obin-chemo+obin arm instead of 0.113%. The latter is the correct value, as shown in Table 5.8 of the ERG report. However, correcting this error had no impact on the CS base-case results since modelling PFS mortality separately per treatment arms was not assumed. It had an impact on scenario analysis 5, as explained in Section 5.2.11, and on the ERG base-case.

b. Correcting adverse event costs for anaemia: As explained in the critique of Section 5.2.9, the ERG found a different value for the unit costs for the management of anaemia (£3,021 instead of £2,117). The ERG updated this value in its base-case model. Changing this parameter had a minor impact on the CS base-case results.
c. Correcting the calculation of the frequency of AEs when AE-related disutilities were incorporated in the cost effectiveness analysis: The ERG noticed that in the calculation of the frequencies of the AEs included in the analysis, wrong Excel cell references were used in the formulas. For instance, the number of events observed for nail infection was used while calculating the frequency of neutropenia. The ERG corrected this error in its base-case model. Correcting this error had no impact on the CS base-case results since AE-related disutilities were not assumed. It had an impact on the CS sensitivity analysis and on the ERG base-case.
d. Correcting the wrong implementation of the "no vial sharing" costs for the obinutuzumab drug acquisition costs: In the calculation of the drug acquisition costs for obinutuzumab, it was assumed that 100%, perfect vial sharing was always attained, even when the "no vial sharing" option was selected in the model. Therefore, the “no vial sharing” assumption was applied to the rituximab arm only. The ERG corrected this error in its base-case model. Correcting this error had no impact on the CS base-case results since vial sharing was assumed. It had an impact on the CS sensitivity analysis and on the ERG base-case.
e. Correcting the errors in sensitivity analyses: The ERG noticed that the upper values used for the administration cost items in the one-way sensitivity analysis were wrong, the savings due to the subcutaneous use of rituximab was not consistently applied and wrong upper and lower values for the subsequent administration costs were used. In the probabilistic sensitivity analysis, the ERG realized that the standard error considered for the parameter "second line treatment costs" was wrong. The company used 20% as the value for the standard error instead of applying a 20% deviation from the mean. The ERG corrected these errors in its base-case model.

Fixing violations

2. Demographic characteristics:
    a. Increased age at baseline: As explained in the ERG critique of Section 5.2.3, the ERG considers that the GALLIUM cohort is younger than the average UK patient. Thus, a model with increased age at baseline is expected to better reflect the UK population. For the ERG’s preferred base-case a baseline age of 62.6 years was chosen as proposed by the company in scenario analysis 3 from Section 5.2.11.
    b. Distribution per chemotherapy regimen: As mentioned in the ERG critique of Section 5.2.3, the appropriate representative average use of the three chemotherapy regimens has some uncertainty. In the ERG preferred base-case, the proportion of UK patients in the GALLIUM trial was considered; thus, 68% bendamustine, 31% CVP and 1% CHOP.
    c. Proportion of females in the population: The ERG considers that gender-averaged values should be based on the proportion of females observed in the UK population. When calculating age and gender specific decrements utilities, the proportion used in the

company's model was 50%. For life-tables, the value used was 50.6%. While this is described as the proportion in the GALLIUM trial in the model and the clarification response, the ERG believes that this might be a reporting error, since this does not match with the 53.2% of females (vs. 46.8% of males) reported in Table 4.3 of this report. The latter values were used in the ERG's preferred base-case analysis.

3. Different PFS and PPS (early PD only) mortality rates for the treatment arms. As mentioned in the critique of Section 5.2.6, the ERG considers that when a treatment effect is sought; different mortality rates for the arms should be used, regardless of statistical significance. The company implemented this option in the updated version of the model and it was chosen by the ERG for its preferred bas-case analysis.
4. Utility decrement with age. As explained in the critique of Section 5.2.8, the ERG does not agree with the company's assumption of not adjusting the utility values for a decline in age. Therefore, the decline in age for utilities was included in the ERG's preferred base-case analysis.

Matters of judgement

5. Use of PFS-IRC data and choice of a Weibull distribution for PFS data extrapolation. In the critique of Section 5.2.6, it is explained that independent review committee assessed PFS data should have been used for the base-case analysis. Furthermore, after reassessment of the goodness of fit for the PFS-IRC data, the ERG chose the Weibull distribution to model PFS in its preferred base-case analysis.
6. Disutilities due to adverse events. As explained in the critique of Section 5.2.8, the ERG included AE disutilities in its preferred base-case as an indirect way to reflect the difference in utilities between the two treatment arms.
7. Vial sharing: In the company's base-case, vial sharing was assumed. This implies that if, for instance 90% of a vial was used in a cycle, only 90% of the costs of that vial was taken into account in the calculations and the remaining unused obinutuzumab could be used in another patient. The ERG considers this assumption less plausible than "no vial sharing", which incurs the full vial cost of each vial opened, as the possibility of using remnant obinutuzumab from an opened vial for another patient may not be always feasible.
8. Adverse event costs and disutilities: As discussed in the critique of Section 5.2.7, the company applied an arbitrary threshold of 2% to the serious adverse events to create a list of the most relevant/frequent ones from the trial. This threshold was applied separately for grade 3, grade 4 and grade 5 adverse events, which led to a situation that grade 3 of an adverse event might be in the list (e.g. pnemonia), whereas grade 4/5 of the same type of adverse event might not. Due to data and time limitations, the ERG did not apply this threshold to the pooled grade 3/4/5 adverse events to create a new list of adverse events to be included in the model. Instead of that, for each adverse event type that was listed in the company submission, the cost and disutilities for all grades higher than or equal to three were incorporated into the cost and QALY calculations of the electronic model (e.g. pnemonia grade 3/4/5 were included in the ERG's analyses even though, based on the 2% threshold rule, only pnemonia grade 3 was included in the company's analyses).

Additional scenarios

The ERG conducted several additional scenario analyses where the structural uncertainties in the ERG preferred base-case were explored. The additional scenario analyses conducted by the ERG are listed below.

Scenario 1. Alternative assumptions on clinical effectiveness

- Scenario 1a. Exploring the impact of the duration of the treatment effect. The ERG performed a threshold analysis to determine at which value of the duration of the treatment effect the resulting ICER was above the £30,000 threshold. Additionally, the ERG included the duration of the treatment effect on the PSA to illustrate how the uncertainty around this parameter might affect the cost effectiveness results.
- Scenario 1b. Assuming a Gompertz distribution to extrapolate PFS data.
- Scenario 1c. Using PFS-INV data (and a Weibull distribution) to model PFS.
- Scenario 1d. Using pooled mortality rates for PFS and early PD health states instead of separately per treatment arm.

Scenario 2. Utilities

- Scenario 2a. Disutilities due to adverse events were not included in the analysis.
- Scenario 2b. Utility decrement with age was not included in the analysis.
- Scenario 2c. Using the utilities reported in the GALLIUM trial for the PFS and PD health states.
- Scenario 2d. Using the utilities from Wild et al.[36] for the PFS and PD health states.
- Scenario 2e. Using the utilities from Bec et al.[44] for the PFS and PD health states.
- Scenario 2f. Using the utilities reported in the GADOLIN trial for PFS and PD health states.
- Scenario 2g. Using the utilities reported in the GALLIUM trial for the PFS health state and mapping FACT-Lym values from Cheung et al.[58] for the PD health state.
- Scenario 2h. Using the utilities reported in the GALLIUM trial for the PFS health state and the utilities reported in the GADOLIN trial for the PD health state.
- Scenario 2i. Using the utilities reported in the GALLIUM trial for the PFS health state and the utilities from Bec et al. for the PD health state.
- Scenario 2j. Considering different utility values for the early and late PD health states. The ERG implemented this as follows. The average of early and late PD utilities from GALLIUM was taken as reference. The ratios of (early PD utility/average utility) and (late PD utility /average utility) were then multiplied by the Wild et al. utility for the PD health state.

Scenario 3. Demographic characteristics

- Scenario 3a. The demographic characteristics from the GALLIUM trial were considered in this scenario.

Scenario 4. Costs

- Scenario 4a. Assuming a chemotherapy distribution according to the UK market research sample (29% bendamustine, 13% CHOP and 36% CVP).
- Scenario 4b. Assuming a chemotherapy distribution according to all patients in the GALLIUM trial (57% bendamustine, 33% CHOP and 10% CVP).
- Scenario 4c. Assuming that 100% of the patients received bendamustine as chemotherapy.
- Scenario 4d. Assuming that 100% of the patients received CHOP as chemotherapy.
- Scenario 4e. Assuming that 100% of the patients received CVP as chemotherapy.
- Scenario 4f. Vial sharing was assumed in this scenario.

5.3.2. Results from the ERG preferred base-case and probabilistic sensitivity analysis

In the ERG base-case analysis, obin-chemo+obin resulted in [redacted] total (discounted) costs and 9.12 total QALYs, while R-chemo+R resulted in [redacted] total (discounted) costs and 8.58 total QALYs, as

presented in Table 5.30 below. Therefore, obin-chemo+obin produced 0.53 additional QALYs at an incremental cost of [redacted] when compared to R-chemo+R, leading to an ICER of [redacted]. This ICER is [redacted] higher than the company base-case ICER of [redacted].

Table 5.30: The ERG preferred base-case deterministic results

Technologies	Total Costs (£)	Total LYG	Tot QALYs	Inc Costs (£)	Inc LYG	Inc QALYs	ICER (£)
obin-chemo+obin	[redacted]	12.79	9.12	[redacted]	0.58	0.53	[redacted]
R-chemo+R	[redacted]	12.22	8.58				
ICER = incremental cost effectiveness ratio; LYG = life years gained; QALYs = quality adjusted life years.							

Disaggregated results for QALYs and costs per health state are presented in Table 5.31 and Table 5.32, respectively. Compared to the company's base-case results, patients spend less time in PFS (73% vs. 79%) and early PD (7% vs. 10%) and more time in late PD (20% vs. 11%) in the ERG base-case analysis.

Table 5.31: Summary of discounted QALY gain by health state

	obin-chemo+obin	R-chemo+R	Difference	Absolute	% of absolute
Health state					
Progression free survival	6.71	5.86	0.85	0.85	73%
Progression < 2 yrs	0.24	0.33	-0.08	0.08	7%
Progression > 2 yrs	2.17	2.40	-0.23	0.23	20%
Total	9.12	8.59	0.53	1.17	100%
Values in the table are discounted and half cycle corrected.					

Disaggregated results for costs are similar to the company's based-case results. This was expected since most of the assumptions considered in the ERG base-case were efficacy- and utility-related, except for the distribution of patients per chemotherapy group, which was favourable to the obin-chemo+obin arm.

Table 5.32: Summary of predicted resource use by category of cost

State	Cost (Obin-chemo)	Cost (R-chemo)	Cost difference	Absolute difference	% of absolute
PFS					
Obinutuzumab	[redacted]	0	[redacted]	[redacted]	[redacted]
Rituximab	0	[redacted]	[redacted]	[redacted]	[redacted]
Chemotherapy	411	406	5	5	[redacted]
Drug Administration	7,760	6,426	1,334	1,334	[redacted]
Adverse Events	737	576	161	161	[redacted]
Supportive Care	7,595	6,807	788	788	[redacted]
PFS Total	[redacted]	[redacted]	[redacted]		
Progressive disease					
Supportive care and subsequent treatment costs	9,762	11,455	-1,693	1,693	[redacted]
Total PD & PFS	[redacted]	[redacted]	[redacted]	[redacted]	100%
Values in the table are discounted and half cycle corrected. PD = progressed disease; PFS = progression free survival					

The ERG performed a PSA on its preferred base-case to explore the uncertainty around the ERG base-case input parameters. Since the ERG corrected some PSA-related parameters, higher uncertainty was expected in the ERG base-case, reflected in wider confidence ellipses for the PSA results and lower cost effectiveness probabilities for the obin-chemo+obin arm. The PSA results, based on 1,000 model iterations as in the CS base-case, are summarised in Table 5.33. These are comparable to the deterministic base-case results presented in Table 5.17 above.

Table 5.33: PSA results of the ERG preferred base-case

Technologies	Total Costs (£)	Total LYG	Tot QALYs	Inc Costs (£)	Inc LYG	Inc QALYs	ICER (£)
obin-chemo+obin	[redacted]	12.71	9.07	[redacted]	0.56	0.52	[redacted]
R-chemo+R	[redacted]	12.15	8.55				
ICER = incremental cost effectiveness ratio; LYG = life years gained; QALYs = quality adjusted life years.							

From the scatterplot of the PSA outcomes on the CE-pane in Figure 5.17, it can be seen that, as expected, the PSA outcomes are more shifted to the left and more scattered, resulting in a few PSA outcomes on the north-western quadrant of the CE plane. This increased uncertainty is also reflected in the CEAC shown in Figure 5.18, where it can be seen that the probability that obin-chemo+obin is cost effective compared to R-chemo+R is approximately [redacted] at a £30,000 per QALY gained threshold. Thus, an absolute reduction of [redacted] in the cost effectiveness probability compared to the company base-case probability at the same threshold.

Figure 5.17: Incremental cost effectiveness scatterplot of the ERG preferred base-case analysis

Figure 5.18: Cost effectiveness acceptability curve of the ERG preferred base-case

5.3.3. Results from the ERG additional exploratory scenario analyses

The additional scenarios listed in Section 5.3.1 were performed on the ERG preferred base-case. The results of these additional scenarios, with the exception of Scenario 1a (which is widely described below), are summarised in Table 5.34.

Table 5.34: Results from the additional scenario analyses conducted on the ERG preferred base-case

Scenarios	Obin-chemo+obin		R-chemo+R		Inc Costs (£)	Inc QALYs	ICER (£)
	Total Costs (£)	Total QALYs	Total Costs (£)	Total QALYs			
CS base-case	[redacted]	10.01	[redacted]	9.23	[redacted]	0.78	[redacted]
ERG preferred base-case	[redacted]	9.12	[redacted]	8.58	[redacted]	0.53	[redacted]
Scenario 1a - (treatment effect duration 5 years)	[redacted]	8.97	[redacted]	8.58	[redacted]	0.39	[redacted]
Scenario 1b - (PFS Gompertz distribution)	[redacted]	9.04	[redacted]	8.52	[redacted]	0.52	[redacted]
Scenario 1c - (PFS-INV data)	[redacted]	8.85	[redacted]	8.24	[redacted]	0.61	[redacted]
Scenario 1d - (pooled mortality)	[redacted]	9.13	[redacted]	8.57	[redacted]	0.56	[redacted]
Scenario 2a - (no AE disutility)	[redacted]	9.12	[redacted]	8.59	[redacted]	0.53	[redacted]
Scenario 2b - (no utility decrement with age)	[redacted]	9.62	[redacted]	9.05	[redacted]	0.58	[redacted]
Scenario 2c - (GALLIUM utilities for PFS and PD health states)	[redacted]	9.87	[redacted]	9.43	[redacted]	0.44	[redacted]
Scenario 2d - (Wild et al. utilities for PFS and PD health states)	[redacted]	8.99	[redacted]	8.47	[redacted]	0.52	[redacted]
Scenario 2e - (Bec et al. utilities for PFS and PD health states)	[redacted]	7.79	[redacted]	7.32	[redacted]	0.47	[redacted]
Scenario 2f - (GADOLIN utilities for PFS and PD health states)	[redacted]	9.60	[redacted]	9.15	[redacted]	0.45	[redacted]
Scenario 2g - (GALLIUM utilities for PFS and mapping FACT-Lym for PD)	[redacted]	9.55	[redacted]	9.08	[redacted]	0.47	[redacted]

Scenario 2h - (GALLIUM utilities for PFS and GADOLIN for PD)	[redacted]	9.67	[redacted]	9.21	[redacted]	0.46	[redacted]
Scenario 2i - (GALLIUM utilities for PFS and Bec et al. for PD)	[redacted]	8.69	[redacted]	8.11	[redacted]	0.59	[redacted]
Scenario 2j - (Different utilities for early and late PD)	[redacted]	9.16	[redacted]	8.63	[redacted]	0.53	[redacted]
Scenario 3a - (demographic characteristics in the GALLIUM trial)	[redacted]	9.43	[redacted]	8.89	[redacted]	0.54	[redacted]
Scenario 4a - (chemotherapy distribution UK market research)	[redacted]	9.12	[redacted]	8.58	[redacted]	0.53	[redacted]
Scenario 4b - (chemotherapy distribution in the GALLIUM trial – all patients)	[redacted]	9.12	[redacted]	8.58	[redacted]	0.53	[redacted]
Scenario 4c - (chemotherapy distribution 100% bendamustine)	[redacted]	9.12	[redacted]	8.58	[redacted]	0.53	[redacted]
Scenario 4d - (chemotherapy distribution 100% CHOP)	[redacted]	9.12	[redacted]	8.58	[redacted]	0.53	[redacted]
Scenario 4e - (chemotherapy distribution 100% CVP)	[redacted]	9.12	[redacted]	8.58	[redacted]	0.53	[redacted]
Scenario 4f - (vial sharing)	[redacted]	9.12	[redacted]	8.58	[redacted]	0.53	[redacted]

AE = adverse event; CHOP = Cyclophosphamide, doxorubicin, vincristine and prednisone; CS = company submission; CVP = Cyclophosphamide, vincristine, and prednisone; ERG = evidence review group; ICER = incremental cost effectiveness ratio; INV = local investigator; PD = progressed disease; PFS = progression free survival; QALYs = quality adjusted life years.

In scenario 1a, the ERG performed a threshold analysis on the duration of the treatment effect. The ICERs obtained when the duration of the treatment effect was varied from 1 to 30 years can be seen in Figure 5.19. Note that the ICER decreases as the duration of the treatment effect increases but this decrease is not linear. Thus, assuming short durations for the treatment effect resulted in very high ICERs, while considering high durations did not change the ICERs so much. In particular, assuming a treatment effect duration of two years, which is the maximum time on maintenance assumed in the model, resulted in an ICER of [redacted]; assuming five years, which is the longest follow-up in the GALLIUM trial and presented as scenario analysis in the CS, resulted in an ICER of [redacted]; and assuming nine years, which is the duration assumed in the CS and ERG base-case, resulted in an ICER of [redacted]. A treatment effect duration of five years was the maximum value assumed where the ICER was above the £30,000 threshold. For longer durations of the treatment effect, the ICER remained below the £30,000 threshold, ranging from [redacted] (six years) to [redacted] (no upper limit). In fact, it was observed that assuming more than 17 years for the treatment effect duration decreased the ICER in approximately [redacted] and the maximum ICER was obtained for 27 years, after that, increasing the treatment effect duration did not change the ICER.

Figure 5.19: Incremental cost effectiveness ratio for different values of the treatment effect duration

Additionally, the ERG performed an additional scenario where the duration of the treatment effect was included in the PSA. In particular, the treatment effect duration was modelled as a uniform distribution between 0 and 18 years. The limits were chosen in such a way that the base-case value of nine years was the expected value of the probability distribution. The lower limit represents then "no treatment effect" and the upper limit can be seen as a very close approximation to "no upper limit for the duration of the treatment effect", given the results of the threshold analysis conducted by the ERG. Nevertheless, it should be noted that the purpose of this scenario was to illustrate how the uncertainty around this parameter affected the cost effectiveness results since, as explained in the critique of Section 5.2.6, the ERG considers that it is uncertain how the value for the treatment effect duration was chosen. The choice of the uniform distribution should be then regarded as exploratory and not based on evidence. This scenario resulted in a probabilistic ICER ([redacted]) which was [redacted] higher than the base-case

probabilistic ICER ([redacted]). While this could be produced by random sampling, it might also be the case that it is a reflection of the non-linear decrease of the ICER as a function of the duration of the treatment effect shown in Figure 5.19 above. As expected, the uncertainty associated to this scenario was increased when compared to the ERG base-case, as reflected in wider confidence ellipses and more PSA outcomes in the north-western quadrant of the CE plane. As a consequence, the cost effectiveness probability for obin-chemo+obin was decreased. In particular, at the £30,000 threshold ICER the probability that obin-chemo+obin is cost effective was approximately [redacted]. Thus, an absolute reduction of [redacted] and [redacted] in the cost effectiveness probability compared to the ERG and company base-case probability at the same threshold, respectively. The scatterplot of the PSA outcomes on the CE-pane and the CEAC for this scenario can be seen in Figure 5.20 and Figure 5.21 below.

Figure 5.20: Incremental cost effectiveness scatterplot of the ERG scenario analysis including treatment effect duration on the PSA

Figure 5.21: Cost effectiveness acceptability curve of the ERG scenario analysis including treatment effect duration on the PSA

In the scenario where a Gompertz distribution was chosen to model progression free survival the ICER was [redacted]. Assuming PFS-INV data to model PFS and a Weibull distribution to extrapolate resulted in an ICER of [redacted]. Finally, when a pooled mortality for both treatment arms was assumed, the ICER obtained was [redacted].

Within the set of scenarios performed on utilities, the ICER ranged from [redacted], when the utilities collected in the GALLIUM trial were used for the PFS health state and the utilities reported in Bec et al. were used for the PD health state, to [redacted], when GALLIUM utilities were used for both PFS and PD health states. This showed that the ICER is sensitive to changes in utilities. In general, it was observed that assuming higher utility values for the PD health state resulted in a higher ICER.

In another scenario, the baseline age was decreased to the one observed in the GALLIUM trial. The resulting ICER was [redacted]. It seems then that the ICER from the ERG preferred base-case is less sensitive to changes in baseline age than the ICER form the company base-case. This may be explained by the inclusion of utility decrement with age in the ERG preferred scenario.

From the results of the cost-related scenarios, it seems that the ICER might be sensitive to changes in the distribution of patients per chemotherapy group, being CHOP and bendamustine the more expensive options and CVP the least expensive. In the hypothetical situation where all patients were assigned to just one chemotherapy group, the ICER obtained were [redacted] for bendamustine, [redacted] for CHOP and [redacted] for CVP.

The ERG has drawn attention to a number of parameters for which it is believed that there is uncertainty and therefore, they should have been included in the probabilistic analyses. This is explained in detail in the critique of Section 5.2.11. Most of these parameters could not be included due to lack of data and time constraints. While this is expected to have a minor/moderate impact on the ICER, the current probabilistic results are likely to underestimate the uncertainty around the model results. However, the ERG considers it difficult to estimate to what extent this is underestimated.

In conclusion, the ERG base-case analysis resulted in an ICER of [redacted] per QALY gained. This ICER value is lower than the £30,000 per QALY threshold. Although the ICER seems to be robust to most of the structural changes explored by the ERG, it is possible that different choices for the treatment effect duration and the utilities for the PD health state would result in an ICER above the £30,000 per QALY threshold.

5.4 Conclusions of the cost effectiveness section

The cost effectiveness searches in the CS were well documented, structured and reproducible. In addition, they were in line with the NICE guide to the methods of technology appraisal.[11]

The ERG considers that the economic model presented in the CS meets the NICE reference case to a reasonable extent. Deviations might have occurred regarding the measurement and valuation of HRQoL but this remained unclear. The economic model is in line with the decision problem formulated by the company but it is only partially in line with the scope. The intervention in the scope is described as "*obinutuzumab in combination with chemotherapy, with or without obinutuzumab maintenance therapy*".[35] However, the company did not include obinutuzumab induction therapy without maintenance therapy in the economic analysis. The comparator included in the company's economic analysis was a rituximab-based chemotherapy, with rituximab maintenance treatment. This comparator is also in line with the scope. However, other relevant comparators listed in the NICE scope (obinutuzumab versus rituximab mono-therapy and bendamustine mono-therapy) were not included in the company's cost effectiveness analysis.

The assessment of the ERG concluded that the model developed by the company was properly described and reported. The company's de novo economic model assessed the cost effectiveness of obinutuzumab in combination with chemotherapy (CVP, CHOP or bendamustine), followed by obinutuzumab maintenance therapy in patients achieving a response, compared to rituximab-based chemotherapy, with rituximab maintenance treatment, in people with untreated advanced follicular lymphoma. This was a four-state cohort transition (Markov) model with monthly cycles and a time horizon of 40 years (changed to 50 years after clarification). The health states considered in the model were progression free (PFS) (on/off treatment), two progressed disease (PD) states, early PD and late PD, and death.

Clinical parameters for the model were derived from the GALLIUM trial data when these were considered mature enough to provide robust estimates. Thus, GALLIUM data were used to estimate time to treatment discontinuation (TTTD), PFS and post progression survival (PPS) for early progressed disease. The investigator (INV) assessed PFS data (PFS-INV) were used, corresponding to the primary endpoint. The extrapolation beyond the observed period in the GALLIUM trial was based on parametric functions. The latest available data cut of GALLIUM with a clinical cut-off date of 10 September 2016 was used. External data were used to populate PPS for late progressed disease using long-term data from the PRIMA trial. The model uses EQ-5D utilities collected from the GALLIUM trial for the PFS health state. Since long-term EQ-5D utility scores collected in the GALLIUM trial were considered immature by the company, post progression utilities were sourced from the literature. Health state related costs consisted of medication costs (induction and maintenance), supportive care costs, subsequent treatment costs in PD, and adverse event costs. Relevant medication costs included costs of obinutuzumab, bendamustine, CHOP, CVP, and rituximab. Resource use was derived from UK reference costs.

The results of the company's base-case cost effectiveness analysis showed that obin-chemo+obin resulted in a total cost of [redacted] and 10.01 QALYs. The comparator, R-chemo+R, resulted in a total cost of [redacted] and 9.23 QALYs. Thus, obin-chemo+obin produced 0.78 additional QALYs at an

incremental cost of [redacted] when compared to R-chemo+R, leading to an ICER of [redacted]. The probabilistic sensitivity analysis results presented by the company estimated that the probability that obin-chemo+obin is cost effective when compared to R-chemo+R is approximately [redacted] at a threshold of £30,000 per QALY gained. The results of different sensitivity and scenario analyses showed that the ICERs remained below [redacted] and were close to the base-case value in most cases. The most influential parameter was the duration of the treatment effect, whose variation resulted in a wide range of possible values of the ICER.

The ERG major concern with respect to the company submission was the validity of some assumptions regarding the implementation of the treatment effectiveness in the economic model. In particular, the duration of the treatment effect was shown to have a major impact on the ICER. Other concerns of the ERG were related to the choice of PFS data (local investigator or independent review committee), the choice of the PFS probability distribution, modelling the same mortality rates for both treatment arms, the generalisability to UK clinical practice (in particular the demographic characteristics of the patient population and the proportion of patients per chemotherapy method) and the estimation of utility and cost input parameters. However, these were shown to have a moderate to minor impact on the cost effectiveness results.

The ERG made various adjustments to the company base-case. These adjustments would have ideally included 1) exploring the impact of the underlying chemotherapy regimen on life years and QALYs gained, 2) using different utilities per treatment arm, 3) correcting inconsistencies between the cost and utility implications of some of the included adverse events (e.g. in the current version of the model febrile neutropenia has only cost implications but no disutility implications), 4) incorporating the costs and disutilities for all adverse events of grade 3 or higher, 5) including the actual/planned dose for obinutuzumab in the obin-CVP subgroup (currently it is not provided), 6) correcting the actual/planned dose for bendamustine (90 mg/m^2 vs. 120 mg/m^2 – there seems to be an inconsistency regarding this value; the ERG does not know whether it is only a reporting error or an actual error in the calculations because the value for bendamustine dosage is hardcoded) and 7) including additional parameters in the probabilistic sensitivity analysis. However, these adjustments were not included due to lack of data and time constraints.

The ERG preferred base-case resulted in an ICER of [redacted] per QALY gained. The ERG's most influential adjustments were 1) choosing of PFS-IRC data and a Weibull distribution for PFS extrapolation; 2) applying a utility decrement by age; 3) increasing age at baseline and 4) considering different mortality rates per treatment arm. From the PSA results, the probability that obin-chemo+obin is cost effective compared to R-chemo+R is approximately [redacted] at a £30,000 per QALY gained threshold. Thus, an absolute reduction of [redacted] in the cost effectiveness probability compared to the company base-case probability at the same threshold.

The ERG conducted several additional scenario analyses where the structural uncertainties in the ERG preferred base-case were explored. The additional scenario analyses conducted by the ERG were categorised in four groups: clinical effectiveness, utilities, demographic characteristics and costs.

The ERG performed a threshold analysis on the duration of the treatment effect. In particular, assuming a treatment effect duration of two years (the maximum time on maintenance assumed in the model) resulted in an ICER of [redacted]. Assuming five years, which is the longest follow-up in the GALLIUM trial and presented as scenario analysis in the company submission, resulted in an ICER of [redacted]. Moreover, a treatment effect duration of [redacted] was the maximum value assumed where the ICER was above the £30,000 threshold. Additionally, the ERG performed another scenario where the duration of the treatment effect was included in the PSA. This was modelled as a uniform distribution between

0 and 18 years. This scenario resulted in a probabilistic ICER ([redacted]) which was [redacted] higher than the base-case probabilistic ICER ([redacted]). The uncertainty associated to this scenario was increased when compared to the ERG base-case, and at the £30,000 threshold ICER the probability that obin-chemo+obin is cost effective was approximately [redacted]. Thus, an absolute reduction of [redacted] and [redacted] in the cost effectiveness probability compared to the ERG and company base-case probability at the same threshold, respectively. In the scenario where a Gompertz distribution was chosen to model progression free survival the ICER was [redacted]. Assuming PFS-INV data to model PFS and a Weibull distribution to extrapolate resulted in an ICER of [redacted]. Finally, when a pooled mortality for both treatment arms was assumed, the ICER obtained was [redacted].

Within the set of scenarios performed on utilities, the ICER ranged from [redacted], when the utilities collected in the GALLIUM trial were used for the PFS health state and the utilities reported in Bec et al. were used for the PD health state, to [redacted], when GALLIUM utilities were used for both PFS and PD health states. This showed that the ICER is sensitive to changes in utilities. In general, it was observed that assuming higher utility values for the PD health state resulted in a higher ICER.

The additional scenario analyses on demographic characteristics and costs showed a moderate to minor impact on the ICER.

Throughout this report, the ERG has drawn attention to a number of parameters for which it is believed that there is uncertainty and therefore, they should have been included in the probabilistic analyses. However, the ERG could not include all of them in the PSA due to lack of data and time constraints. This is explained in detail in the critique of Section 5.2.11. While including all these parameters in the probabilistic analyses of the model is expected to have a minor/moderate impact on the ICER, the current probabilistic results are likely to underestimate the uncertainty around the model results. However, the ERG considers it difficult to estimate to what extent the uncertainty is underestimated.

In conclusion, the ERG base-case analysis resulted in an ICER of [redacted] per QALY gained. This ICER value is lower than the £30,000 per QALY threshold. Although the ICER seems to be robust to most of the structural changes explored by the ERG, it is possible that different choices for the treatment effect duration and the utilities for the PD health state would result in an ICER above the £30,000 per QALY threshold.

6. IMPACT ON THE ICER OF ADDITIONAL CLINICAL AND ECONOMIC ANALYSES UNDERTAKEN BY THE ERG

The ERG preferred base-case was presented in Section 5.3, which was the result of applying various changes to the company's base-case. In Table 6.1 can be seen how each individual change affects the ICER; the last row shows the combined effect of all changes simultaneously, which corresponds to the ERG preferred base-case reported in Section 5.3.

Table 6.1: Revised base case cost effectiveness analysis, incorporating corrections and amendments identified by the ERG

Scenarios	Obin-chemo+obin		R-chemo+R		Inc Costs (£)	Inc QALYs	ICER (£)
	Total Costs (£)	Total QALYs	Total Costs (£)	Total QALYs			
0. CS base-case	[redacted]	10.01	[redacted]	9.23	[redacted]	0.78	[redacted]
1. Fixing errors	[redacted]	10.01	[redacted]	9.23	[redacted]	0.78	[redacted]
(1+2a). Fixing errors and increasing baseline age	[redacted]	9.63	[redacted]	8.92	[redacted]	0.71	[redacted]
(1+2b). Fixing errors and chemotherapy distribution	[redacted]	10.01	[redacted]	9.23	[redacted]	0.78	[redacted]
(1+2c). Fixing errors and female distribution	[redacted]	10.02	[redacted]	9.24	[redacted]	0.79	[redacted]
(1+3). Fixing errors and different mortality per arm	[redacted]	9.98	[redacted]	9.26	[redacted]	0.72	[redacted]
(1+4). Fixing errors and age utility decrement	[redacted]	9.48	[redacted]	8.76	[redacted]	0.72	[redacted]
(1+5). Fixing errors and PFS-IRC Weibull	[redacted]	10.00	[redacted]	9.34	[redacted]	0.66	[redacted]
(1+6). Fixing errors and AE disutility	[redacted]	10.01	[redacted]	9.23	[redacted]	0.78	[redacted]
(1+7). Fixing errors and no vial sharing	[redacted]	10.01	[redacted]	9.23	[redacted]	0.78	[redacted]
(1+8). Fixing errors and AE grade ≥3 costs & disutilities	[redacted]	10.01	[redacted]	9.23	[redacted]	0.78	[redacted]
(1 to 8 all): ERG preferred base-case	[redacted]	9.12	[redacted]	8.58	[redacted]	0.53	[redacted]

AE = adverse event; CS = company submission; ICER = incremental cost effectiveness ratio; IRC = independent review committee; PFS = progression free survival; QALYs = quality adjusted life years.

7. OVERALL CONCLUSIONS

7.1 Statement of principal findings

The CS includes a systematic review of the RCT evidence for obinutuzumab in previously untreated FL. One RCT was identified (GALLIUM) that investigated the efficacy and safety of obinutuzumab with chemotherapy followed by obinutuzumab maintenance for responders. This was a large, well-conducted randomised trial including 1,202 patients with FL. The trial was conducted at 177 centres in 18 countries (293 (21%) patients were from the UK and almost 50% were from Western Europe).

The primary outcome was investigator-assessed PFS. Key secondary outcomes included PFS assessed by independent review committee (IRC), overall survival and response rates. Health-related quality of life was also assessed using a disease-specific tool (FACT-Lym) and EQ-5D. The submission focused on results of effectiveness on a data cut of January 2016. On request the company provided full results for the later cut-off of September 2016. We have provided results for both time points in this report.

Although GALLIUM is a good quality RCT, a number of limitations were identified by the ERG. The trial was open-label, therefore results based on independent review will be less prone to bias than investigator results. In the trial, obinutuzumab and its comparator rituximab could be given to patients with three different chemotherapy regimens (CHOP, CVP and bendamustine). In the trial approximately 57% received bendamustine, 33% CHOP and 10% CVP. The breakdown of the chemotherapy used may not be reflective of the UK. The trial was not designed to investigate differences in chemotherapy regimens so any variation in results between chemotherapy regimens may reflect genuine differences of effectiveness or patient selection factors.

Although GALLIUM had a reasonable follow-up duration, data were not fully mature for the main outcomes. Median progression-free survival (PFS) could not be determined. Overall obinutuzumab was superior to rituximab for PFS (HR = 0.72 (0.56 to 0.93)) for the latest cut-off using IRC data. Although outcomes relating to progression were positive, no differential effects on HRQoL were identified. The committee will need to consider whether improvements in PFS and possible delay to new anti-lymphoma medication are worthwhile alone. The committee will further need to consider any possible relationship between improvements in PFS and improvements in overall survival. Overall survival data in GALLIUM were not mature. GALLIUM is an ongoing trial which should provide, further, more mature results. Finally, the higher rate of serious and higher grade events with obinutuzumab needs to be considered in terms of management of the disease and acceptability to patients.

The results of the company's base-case cost effectiveness analysis showed that obin-chemo+obin resulted in a total cost of [redacted] and 10.01 QALYs. The comparator, R-chemo+R, resulted in a total cost of [redacted] and 9.23 QALYs. Thus, obin-chemo+obin produced 0.78 additional QALYs at an incremental cost of [redacted] when compared to R-chemo+R, leading to an ICER of [redacted]. The probabilistic sensitivity analysis results presented by the company estimated that the probability that obin-chemo+obin is cost effective compared to R-chemo+R is approximately [redacted] at a threshold of £30,000 per QALY gained. The results of different sensitivity and scenario analyses showed that the ICERs remained below [redacted] and were close to the base-case value in most cases. The most influential parameter was the duration of the treatment effect, whose variation resulted in a wide range of possible values of the ICER.

The ERG made various adjustments to the company base-case. These adjustments would have ideally included 1) exploring the impact of the underlying chemotherapy regimen on life years and QALYs gained, 2) using different utilities per treatment arm, 3) correcting inconsistencies between the cost and

utility implications of some of the included adverse events (e.g. in the current version of the model febrile neutropenia has only cost implications but no disutility implications), 4) incorporating the costs and disutilities for all adverse events of grade 3 or higher, 5) including the actual/planned dose for obinutuzumab in the obin-CVP subgroup (currently it is not provided), 6) correcting the actual/planned dose for bendamustine (90 mg/m^2 vs. 120 mg/m^2 – there seems to be an inconsistency regarding this value; the ERG does not know whether it is only a reporting error or an actual error in the calculations because the value for bendamustine dosage is hardcoded) and 7) including additional parameters in the probabilistic sensitivity analysis. However, these adjustments were not included due to lack of data and time constraints.

The ERG preferred base-case resulted in an ICER of [redacted] per QALY gained. The ERG's most influential adjustments were 1) choosing of PFS-IRC data and a Weibull distribution for PFS extrapolation; 2) applying a utility decrement by age; 3) increasing age at baseline and 4) considering different mortality rates per treatment arm. From the PSA results, the probability that obin-chemo+obin is cost effective compared to R-chemo+R is approximately [redacted] at a £30,000 per QALY gained threshold. Thus, an absolute reduction of [redacted] in the cost effectiveness probability compared to the company base-case probability at the same threshold.

The ERG performed a threshold analysis on the duration of the treatment effect. In particular, assuming a treatment effect duration of two years (the maximum time on maintenance assumed in the model) resulted in an ICER of [redacted]. Assuming five years, which is the longest follow-up in the GALLIUM trial and presented as scenario analysis in the company submission, resulted in an ICER of [redacted]. Moreover, a treatment effect duration of [redacted] was the maximum value assumed where the ICER was above the £30,000 threshold. Additionally, the ERG performed an additional scenario where the duration of the treatment effect was included in the PSA. This was modelled as a uniform distribution between 0 and 18 years. This scenario resulted in a probabilistic ICER ([redacted]) which was [redacted] higher than the base-case probabilistic ICER ([redacted]). The uncertainty associated to this scenario was increased when compared to the ERG base-case, and at the £30,000 threshold ICER the probability that obin-chemo+obin is cost effective was approximately [redacted]. Thus, an absolute reduction of [redacted] and [redacted] in the cost effectiveness probability compared to the ERG and company base-case probability at the same threshold, respectively. In the scenario where a Gompertz distribution was chosen to model progression free survival the ICER was [redacted]. Assuming PFS-INV data to model PFS and a Weibull distribution to extrapolate resulted in an ICER of [redacted]. Finally, when a pooled mortality for both treatment arms was assumed, the ICER obtained was [redacted].

Within the set of scenarios performed on utilities, the ICER ranged from [redacted], when the utilities collected in the GALLIUM trial were used for the PFS health state and the utilities reported in Bec et al. were used for the PD health state, to [redacted], when GALLIUM utilities were used for both PFS and PD health states. This showed that the ICER is sensitive to changes in utilities. In general, it was observed that assuming higher utility values for the PD health state resulted in a higher ICER. The additional scenario analyses on demographic characteristics and costs showed a moderate to minor impact on the ICER.

The ERG has drawn attention to a number of parameters for which it is believed that there is uncertainty and therefore, they should have been included in the probabilistic analyses. Most of these parameters could not be included due to lack of data and time constraints. While this is expected to have a minor/moderate impact on the ICER, the current probabilistic results are likely to underestimate the uncertainty around the model results. However, the ERG considers it difficult to estimate to what extent this is underestimated.

In conclusion, the ERG base-case analysis resulted in an ICER of [redacted] per QALY gained. This ICER value is lower than the £30,000 per QALY threshold. Although the ICER seems to be robust to most of the structural changes explored by the ERG, it is possible that different choices for the treatment effect duration and the utilities for the PD health state would result in an ICER above the £30,000 per QALY threshold.

7.2 Strengths and limitations of the assessment

The clinical evidence is based on a good quality randomised controlled trial including 1202 patients with follicular lymphoma. The comparator arm was rituximab, a valid comparator for this appraisal and in clinical practice. Outcomes assessed reflect the scope and are relevant to patients in practice.

Weaknesses include the issue that the breakdown of chemotherapy used with intervention and comparator may not be reflective of UK practice. In the trial, obinutuzumab and its comparator rituximab could be given to patients with three different chemotherapy regimens (CHOP, CVP and bendamustine). Approximately 57% received bendamustine, 33% CHOP and 10% CVP. The breakdown of the chemotherapy used may not be reflective of the UK. The trial was not designed to investigate differences in chemotherapy regimens so any variation in results between chemotherapy regimens may reflect genuine differences of effectiveness or patient selection factors.

Although GALLIUM had a reasonable follow up duration, data were not fully mature for the main outcomes. Median progression-free survival (PFS) could not be determined and overall survival data in GALLIUM were not mature.

The main strength of the CS is that the cost effectiveness section is well structured and the cost effectiveness analyses have been reported transparently. Furthermore, the analyses of the survival data were correctly performed, following the guidance from the NICE Decision Support Unit. For the extrapolation of progression free survival beyond the trial period, parametric functions were fitted simultaneously for both treatment arms data, with treatment as a covariate in the model, which allowed accommodating both proportional hazards and accelerated failure time models. Additionally, the structure of the model developed by the company is in line with other, commonly used, Markov models for progression in oncology but it has the advantage of incorporating early and late progressed disease health states, which seems to be appropriate for the decision problem at hand. The model also includes relevant adverse events, utilities and costs. Sensitivity analyses were performed on the model parameters and the results were robust to most of the structural assumptions.

The main weakness of the cost effectiveness section of the company submission is the reliance on assumptions that could not be verified with the presented evidence. In particular, the duration of the treatment effect and the choice of the progression free survival probability distribution might have a major impact on the cost effectiveness results. The health-related quality of life section of the company submission is sometimes lacking transparency. It remained unclear whether an UK tariff for non-UK EQ-5D utility values was applied or not and whether the effects of adverse events while on treatment were captured in the utility values collected in the GALLIUM trial. Furthermore, despite being unpublished, inconsistent with the results of the GALLIUM trial and unverifiable (by the ERG), the company relied on the utility values reported by Wild et al.[36] to inform the utilities assigned to the progressed disease health state. The choice of these utilities might also have a significant impact on the cost effectiveness results. Finally, it also remained uncertain whether the proportions of patients treated with each chemotherapy method (bendamustine, CHOP, CVP) used in the GALLIUM trial are reflective or not of the UK clinical practice. Since GALLIUM was not powered to detect differences between the three chemotherapy methods and patients were not randomised to chemotherapies, it was

not feasible to conduct a robust scenario analysis where PFS and OS estimates are obtained with a different proportion of chemotherapy regimens. However, if there were any treatment effect due to the underlying chemotherapy method, this would not be possible to detect with the current analyses.

7.3 Suggested research priorities

GALLIUM is an ongoing trial which should provide, further, more mature results. Further research might include an investigation comparing obinutuzumab with different chemotherapy regimens (CHOP, CVP and bendamustine).

The ERG considers that the cost effectiveness analyses would benefit from long-term follow up data (from GALLIUM) which could be used to validate the key assumptions on the duration of the treatment effect and the extrapolation of the PFS parametric curves. Likewise, a more recent measurement of utility values for the progressed disease health state would improve the transparency of the health-related quality of life section of the economic evaluation.

8. REFERENCES

[1] Roche Products Limited. *Single technology appraisal (STA). Gazyvaro (obinutuzumab) in combination with chemotherapy for the first-line treatment of patients with advanced follicular lymphoma: Company evidence submission to National Institute of Health and Care Excellence [ID520]*: Roche Products Limited, May 2017 [accessed 19.5.17]. 216p.

[2] Office for National Statistics. Cancer Registration Statistics, England [Internet]. ONS, April 2017 [accessed April 2017]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/cancerregistrationstatisticscancerregistrationstatisticsengland

[3] Casulo C, Byrtek M, Dawson KL, Zhou X, Farber CM, Flowers CR, et al. Early relapse of follicular lymphoma after rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone defines patients at high risk for death: an analysis from the National LymphoCare Study. *J Clin Oncol* 2015;33(23):2516-22.

[4] Tan D, Horning SJ, Hoppe RT, Levy R, Rosenberg SA, Sigal BM, et al. Improvements in observed and relative survival in follicular grade 1-2 lymphoma during 4 decades: the Stanford University experience. *Blood* 2013;122(6):981-7.

[5] Haematological Malignancy Research Network (HMRN). Survival: Follicular lymphoma [Internet]. York: Haematological Malignancy Research Network (HMRN) n.d. [accessed 19.6.17]. Available from: https://www.hmrn.org/statistics/survival

[6] Cheung MC, Imrie KR, Friedlich J, Buckstein R, Lathia N, Mittmann N. The impact of follicular (FL) and other indolent non-Hodgkin's lymphomas (NHL) on work productivity-a preliminary analysis. *Psychooncology* 2009;18(5):554-9.

[7] National Institute for Health and Care Excellence. *Rituximab for the first-line treatment of stage III-IV follicular lymphoma. NICE technology appraisal guidance (TA243) [Internet]*. London: NICE, 2012 [accessed 20.6.17]. 39p. Available from: https://www.nice.org.uk/guidance/ta243

[8] Roche Products Limited. *Gazyvaro (obinutuzumab) in combination with chemotherapy for the first-line treatment of patients with advanced follicular lymphoma (ID1020): response to request for clarification from the ERG*: Roche Products Limited., 2017 [accessed 26.6.17]. 39p.

[9] Lymphoma Association. Follicular lymphoma [Internet]. Aylesbury: Lymphoma Association, September 2014 [accessed 20.6.17]. Available from: https://www.lymphomas.org.uk/about-lymphoma/types/non-hodgkin-lymphoma/follicular-lymphoma

[10] Roche Products Limited. *Gazyvaro (obinutuzumab) in combination with chemotherapy for the first-line treatment of patients with advanced follicular lymphoma [ID1020]: response to request for clarification from the ERG. Hand searches strategies (v0.1 220617) [Excel spreadsheet provided with the company's response]*: Roche Products Limited., 2017 [accessed 26.6.17]

[11] National Institute for Health and Care Excellence. *Guide to the methods of technology appraisal 2013 [Internet]*. London: NICE, 2013 [accessed 17.7.2017]. 93p. Available from: https://www.nice.org.uk/process/pmg9/chapter/foreword

[12] Roche Products Limited. *Single technology appraisal (STA): Gazyvaro (obinutuzumab) in combination with chemotherapy for the first-line treatment of patients with advanced follicular lymphoma. Supporting Appendices to Company evidence submission [ID520]*: Roche Products Limited, May 2017 [accessed 19.5.17]. 46p.

[13] Roche Products Limited. *Gazyvaro (obinutuzumab) in combination with chemotherapy for the first-line treatment of patients with advanced follicular lymphoma [ID1020]: response to request for clarification from the ERG. Revised electronic search strategies (v0.1 220617) [Excel spreadsheet provided with the company's response]*: Roche Products Limited., 2017 [accessed 26.6.17]

[14] Canadian Agency for Drugs and Technologies in Health. *CADTH peer review checklist for search strategies [Internet]*. Ottawa: CADTH, 2013 [accessed 17.7.13]. 3p. Available from: http://www.cadth.ca/en/resources/finding-evidence-is

[15] National Institute for Health and Care Excellence. *Single Technology Appraisal: company evidence submission template [Internet]*. London: NICE, 2015 [accessed 11.6.15]. 28p. Available from: http://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/specification-for-company-submission-of-evidence-2015-version.docx

[16] Scottish Intercollegiate Guidelines Network. *Search filters: randomised controlled trials [Internet]*. Edinburgh: SIGN, n.d. [accessed 10.7.17]. 3p. Available from: http://www.sign.ac.uk/assets/search-filters-randomised-controlled-trials.docx

[17] Centre for Reviews and Dissemination. *Systematic Reviews: CRD's guidance for undertaking reviews in health care [Internet]*. York: University of York, 2009 [accessed 23.3.11] Available from: http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm

[18] John Wiley and Sons Inc. Cochrane Library. Search Manager Tab help and tips: phrase searching [Internet]. Oxford: Wiley Online Library, n.d. [accessed 10.7.17]. Available from: http://www.cochranelibrary.com/

[19] Cheson BD, Pfistner B, Juweid ME, Gascoyne RD, Specht L, Horning SJ, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol* 2007;25(5):579-86.

[20] F. Hoffmann-La Roche Ltd. *Primary Clinical Study Report – BO21223 – A Multicenter, Phase III, Open-Label, Randomized Study in Previously Untreated Patients with Advanced Indolent Non-Hodgkin's Lymphoma Evaluating the Benefit of GA101 (RO5072759) plus Chemotherapy Compared with Rituximab plus Chemotherapy Followed by GA101 or Rituximab Maintenance Therapy in Responders. Report No. 1067980, September 2016. Protocol BO21223 (Clinical Study Report. Data cut-off: 31-Jan-2016) [PDF provided with the company's submission]*: F. Hoffmann-La Roche Ltd., September 2016 [accessed 19.6.17]. 4973p.

[21] Roche Products Limited. *Gazyvaro (obinutuzumab) in combination with chemotherapy for the first-line treatment of patients with advanced follicular lymphoma [ID1020]: response to request for clarification from the ERG. Assessment of delayed-entry adverse event records and update of safety results reported at the primary analysis in study BO21223 (GALLIUM) (CLINICAL CUTOFF: 31 JANUARY 2016). Obinutuzumab (RO5072759) Assessment report [PDF document provided with the company's response]*: Roche Products Limited., 9th June 2017 [accessed 3.7.17]. 733p.

[22] Grigg A, Dyer MJ, Diaz MG, Dreyling M, Rule S, Lei G, et al. Safety and efficacy of obinutuzumab with CHOP or bendamustine in previously untreated follicular lymphoma. *Haematologica* 2017;102(4):765-772.

[23] Scottish Intercollegiate Guidelines Network. *Search filters: economic studies [Internet]*. Edinburgh: SIGN, n.d. [accessed 10.7.17]. 3p. Available from: http://www.sign.ac.uk/assets/search-filters-economic-studies.docx

[24] Arber M, Garcia S, Veale T, Edwards M, Shaw A, Glanville J. *Sensitivity of a search filter designed to identify studies reporting health state utility values [Internet]*. York: YHEC, 2015 [accessed 26.6.17]. 1p. Available from: http://www.indirect-treatment-comparisons.com/wp-content/uploads/2015/06/Poster-374-Sensitivity-Of-A-Search-Filter.pdf

[25] Lefebvre C, Manheimer E, Glanville J. 6.4.9 Language, date and document format restrictions (Chapter 6: Searching for studies) [Internet]. In: Higgins JPT, Green S, editors. Version 5.1.0 [updated March 2011]: The Cochrane Collaboration, 2011 [accessed 25.5.17]. Available from: http://handbook.cochrane.org/

[26] Dundar Y, Bagust A, Hounsome J, McLeod C, Boland A, Davis H, et al. Rituximab for the first-line treatment of stage III/IV follicular non-Hodgkin's lymphoma. *Health Technol Assess* 2009;13 Suppl 1:23-8.

[27] Ray JA, Carr E, Lewis G, Marcus R. An evaluation of the cost-effectiveness of rituximab in combination with chemotherapy for the first-line treatment of follicular non-Hodgkin's lymphoma in the UK. *Value Health* 2010;13(4):346-57.

[28] Papaioannou D, Rafia R, Rathbone J, Stevenson M, Buckley Woods H, Stevens J. Rituximab for the first-line treatment of stage III-IV follicular lymphoma (review of Technology Appraisal No. 110): a systematic review and economic evaluation. *Health Technol Assess* 2012;16(37):1-253, iii-iv.

[29] Greenhalgh J, Bagust A, Boland A, Blundell M, Oyee J, Beale S, et al. Rituximab for the first-line maintenance treatment of follicular non-Hodgkin's lymphoma : a NICE single technology appraisal. *PharmacoEcon* 2013;31(5):403-13.

[30] Dewilde S, Woods B, Castaigne JG, Parker C, Dunlop W. Bendamustine-rituximab: a cost-utility analysis in first-line treatment of indolent non-Hodgkin's lymphoma in England and Wales. *J Med Econ* 2014;17(2):111-24.

[31] Wang H, Aas E, Smith A, Courch S, Roman E, Patmore R. Forecasting treatment costs of follicular lymphoma: a population-based discrete event simulation (PCN76). *Value Health* 2016;19:A147.

[32] Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ* 1996;313(7052):275-83.

[33] Hoffmann-La Roche, German Low Grade Lymphoma Study Group, Institute of Cancer Research (UK). A study of obinutuzumab (RO5072759) plus chemotherapy in comparison with rituximab plus chemotherapy followed by obinutuzumab or rituximab maintenance in patients with untreated advanced indolent Non-Hodgkin's Lymphoma (GALLIUM). NCT01332968. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2017. Available from: https://ClinicalTrials.gov/show/NCT01332968

[34] Maurer MJ, Bachy E, Ghesquieres H, Ansell SM, Nowakowski GS, Thompson CA, et al. Early event status informs subsequent outcome in newly diagnosed follicular lymphoma. *Am J Hematol* 2016;91(11):1096-1101.

[35] National Institute for Health and Care Excellence. *Single Technology Appraisal. Obinutuzumab for untreated advanced follicular lymphoma: Final scope (Appendix B) [Internet]*. London: NICE, March 2017 [accessed 9.3.17]. 4p. Available from: https://www.nice.org.uk/guidance/gid-ta10137/documents/final-scope

[36] Wild D, Walker M, Pettengell R, Lewis G. Utility elicitation in patients with follicular lymphoma (PCN62). *Value Health* 2006;9:A294.

[37] Haematological Malignancy Research Network. *Patient's age and treatment for haematological malignancy: a report from the Haematological Malignancy Research Network (HMRN) [PDF provided with the company's submission]*. York: HMRN, 2014 [accessed 20.6.17]. 39p.

[38] Wallington M. Variations in body surface area of patients receiving chemotherapy treatment in England: Systemic Anti-Cancer Therapy (SACT) Chemotherapy Dataset [Internet]. Oxford: Chemotherapy Intelligence Unit, 2013 [accessed 4.7.17]: 2.

[39] Salles G, Seymour JF, Offner F, Lopez-Guillermo A, Belada D, Xerri L, et al. Rituximab maintenance for 2 years in patients with high tumour burden follicular lymphoma responding to rituximab plus chemotherapy (PRIMA): a phase 3, randomised controlled trial. *Lancet* 2011;377(9759):42-51.

[40] 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(6):743-54.

[41] Roche Products Limited. *Gazyvaro (obinutuzumab) in combination with chemotherapy for the first-line treatment of patients with advanced follicular lymphoma [ID1020]: response to request for clarification from the ERG. Cost-effectiveness Markov model of Gazyvaro compared to Mabthera regimens in first-line in Follicular lymphoma (Clarification CE model v0.1 220617) [Excel spreadsheet provided with the company's response]*: Roche Products Limited., 2017 [accessed 26.6.17]

[42] Salles GA, Seymour JF, Feugier P, Offner F, Lopez-Guillermo A, Belada D, et al. Updated 6 year follow-up of the PRIMA study confirms the benefit of 2-year rituximab maintenance in follicular lymphoma patients responding to frontline immunochemotherapy. Abstract presented at the 55th Annual Meeting of the American Society of Hematology (ASH); 7-10 Dec 2013; New Orleans: US. *Blood* 2013;122(21):no pagination.

[43] Nastoupil LJ, Sinha R, Byrtek M, Ziemiecki R, Taylor M, Friedberg JW, et al. Comparison of the effectiveness of frontline chemoimmunotherapy regimens for follicular lymphoma used in the United States. *Leuk Lymphoma* 2015;56(5):1295-302.

[44] Bec M, Cognet M, Taieb V, Pacou M, Gauthier A. French utility elicitation in previously treated European patients with indolent Non-Hodgkin Lymphoma (INHL). *Value Health* 2014;17(7):A646.

[45] Ara R, Brazier JE. Using health state utility values from the general population to approximate baselines in decision analytic models when condition-specific data are not available. *Value Health* 2011;14(4):539-45.

[46] Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins J. Health state utilities for non small cell lung cancer. *Health Qual Life Outcomes* 2008;6:84.

[47] National Institute for Health and Care Excellence. *Pixantrone monotherapy for treating multiply relapsed or refractory aggressive non-Hodgkin's B-cell lymphoma. NICE Technology Appraisal Guidance 306 [Internet]*. London: NICE, 2014 [accessed 12.7.17]. 57p. Available from: https://www.nice.org.uk/guidance/ta306

[48] Tolley K, Goad C, Yi Y, Maroudas P, Haiderali A, Thompson G. Utility elicitation study in the UK general public for late-stage chronic lymphocytic leukaemia. *Eur J Health Econ* 2013;14(5):749-59.

[49] Swinburn P, Lloyd A, Nathan P, Choueiri TK, Cella D, Neary MP. Elicitation of health state utilities in metastatic renal cell carcinoma. *Curr Med Res Opin* 2010;26(5):1091-6.

[50] Beusterien KM, Davies J, Leach M, Meiklejohn D, Grinspan JL, O'Toole A, et al. Population preference values for treatment outcomes in chronic lymphocytic leukaemia: a cross-sectional utility study. *Health Qual Life Outcomes* 2010;8:50.

[51] Levy V, Porcher R, Leblond V, Fermand JP, Cazin B, Maloisel F, et al. Evaluating treatment strategies in advanced Waldenstrom macroglobulinemia: use of quality-adjusted survival analysis. *Leukemia* 2001;15(9):1466-70.

[52] Fischbach W, Schramm S, Goebeler E. Outcome and quality of life favour a conservative treatment of patients with primary gastric lymphoma. *Z Gastroenterol* 2011;49(4):430-5.

[53] Korszun A, Sarker SJ, Chowdhury K, Clark C, Greaves P, Johnson R, et al. Psychosocial factors associated with impact of cancer in longterm haematological cancer survivors. *Br J Haematol* 2014;164(6):790-803.

[54] Friedlich J, Cheung M, Imrie K, Hales B, Mittmann N, Buckstein R. Discrimination of health states in follicular lymphoma with utilities derived from the EuroQOL EQ5D instrument. *Blood* 2006;108:3329.

[55] Andrade-Campos MM, Montes-Limon AE, Soro-Alcubierre G, Grasa JM, Lopez-Gomez L, Baringo T, et al. Long-term efficacy of (90)Y ibritumomab tiuxetan therapy in follicular non-Hodgkin lymphoma and health-related quality of life. *Ann Hematol* 2014;93(12):1985-92.

[56] Olin RL, Kanetsky PA, Ten Have TR, Nasta SD, Schuster SJ, Andreadis C. Determinants of the optimal first-line therapy for follicular lymphoma: a decision analysis. *Am J Hematol* 2010;85(4):255-60.

[57] Pettengell R, Donatti C, Hoskin P, Poynton C, Kettle PJ, Hancock B, et al. The impact of follicular lymphoma on health-related quality of life. *Ann Oncol* 2008;19(3):570-576.

[58] Cheung YB, Thumboo J, Gao F, Ng GY, Pang G, Koo WH, et al. Mapping the English and Chinese versions of the Functional Assessment of Cancer Therapy-General to the EQ-5D utility index. *Value Health* 2009;12(2):371-6.

[59] Roche Products Limited. *ID841 Gazyvaro in combination with bendamustine for treating rituximab-refractory follicular lymphoma: response to clarification questions [Internet]*: Roche Products Limited, June 2016 [accessed 24.5.17]. 68p. Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta10020/documents

[60] National Institute for Health and Care Excellence. *Single Technology Appraisal. Obinutuzumab in combination with bendamustine for treating rituximab-refractory follicular lymphoma: Committee Papers (ID841) [Internet]*. London: NICE, 2017 [accessed 25.5.17]. 545p. Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta10020/documents

[61] Dolan P, Roberts J. Modelling valuations for Eq-5d health states: an alternative model using differences in valuations. *Med Care* 2002;40(5):442-6.

[62] Devlin N, Shah K, Feng Y, Mulhern B, Van Hout B. *Valuing Health-Related Quality of Life: an EQ-5D-5L Value Set for England (Research Paper 16/01) [Internet]*. London: Office for Health Economics, 2016 [accessed 12.7.17]. 26p. Available from: https://www.ohe.org/publications/valuing-health-related-quality-life-eq-5d-5l-value-set-england

[63] F. Hoffmann-La Roche Ltd. *Update Clinical Study Report – BO21223 – A Multicenter, Phase III, Open-Label, Randomized Study in Previously Untreated Patients with Advanced Indolent Non-Hodgkin's Lymphoma Evaluating the Benefit of GA101 (RO5072759) plus Chemotherapy Compared with Rituximab plus Chemotherapy Followed by GA101 or Rituximab Maintenance Therapy in Responders. Report No. 1075139, May 2017. Protocol BO21223 (Clinical Study Report. Clinical cutoff date: 10 September 2016) [PDF provided with the company's clarification response]*: F. Hoffmann-La Roche Ltd., May 2017 [accessed 6.7.17]. 8564p.

[64] Roche Products Limited. *Gazyvaro (obinutuzumab) in combination with chemotherapy for the first-line treatment of patients with advanced follicular lymphoma [ID1020]: response to request for clarification from the ERG. Appendix A: Additional scenario analyses and model results for clarification answers*: Roche Products Limited., 2017 [accessed 26.6.17]. 7p.

[65] Lewis G, Marcus R, Proctor S, Gyldmark M, Creeden J, Aultman R, et al. The cost-effectiveness of rituximab, cyclophosphamide, vincristine and prednisolone (RCVP) compared with CVP for the treatment of follicular non-hodgkin's lymphoma (NHL) in the UK. *Blood* 2006;108:345.

[66] Joint Formulary Committee. *British National Formulary [Internet]*. London: BMJ Group and Pharmaceutical Press, 2017 [accessed April 2017]. Available from: https://www.medicinescomplete.com/mc/bnf/current/

[67] Department of Health. Drugs and pharmaceutical electronic market information (eMit) [Internet]. London: Department of Health, 2016 [accessed March 2017]. Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit

[68] South East London Cancer Network. Bendamustine for Indolent Non-Hodgkin's Lymphoma. Version 1 [Internet]. In. London: London Cancer Alliance (LCA), 28th Feb 2013 [accessed 4.7.17]: 2. Available from: http://www.londoncanceralliance.nhs.uk/media/42239/Bendamustine_Relapsed_Indolent_NHL_Protocol_v1.0.pdf

[69] Department of Health. *NHS reference costs 2015-2016 [Internet]*. London: Department of Health, 2017 [accessed April 2017] Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/577083/Reference_Costs_2015-16.pdf

[70] Curtis L, Burns A, Personal Social Services Research Unit, eds. *Unit Costs of Health and Social Care [PDF provided with the company's submission]*. Canterbury: University of Kent, 2016 [accessed 8.5.17]. Available from: http://www.pssru.ac.uk/project-pages/unit-costs/2016/index.php?file=full

[71] Roche. *MabThera (rituximab) for the first-line treatment of follicular non-Hodgkin's lymphoma: submission to the National Institute for Health and Clinical Excellence (NICE)*. London: NICE, 2010

[72] Dreyling M, Ghielmini M, Rule S, Salles G, vitolo U, Ladetto M, et al. Newly diagnosed and relapsed follicular lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2016;27(Suppl 5):v83–v90.

[73] Department of Health. *National schedule of reference costs: the main schedule (Excel spreadsheet) [Internet]*. London: Department of Health, 2016 [accessed 6.7.17] Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/577084/National_schedule_of_reference_costs_-_main_schedule.xlsx

[74] Pavey T, Hoyle M, Ciani O, Crathorne L, Jones-Hughes T, Cooper C, et al. Dasatinib, nilotinib and standard-dose imatinib for the first-line treatment of chronic myeloid leukaemia: systematic reviews and economic analyses. *Health Technol Assess* 2012;16(42):iii-iv, 1-277.

[75] van Oers MH, Van Glabbeke M, Giurgea L, Klasa R, Marcus RE, Wolf M, et al. Rituximab maintenance treatment of relapsed/resistant follicular non-Hodgkin's lymphoma: long-term outcome of the EORTC 20981 phase III randomized intergroup study. *J Clin Oncol* 2010;28(17):2853-8.

[76] Kaltenthaler E, Carroll C, Hill-McManus D, Scope A, Holmes M, Rice S, et al. The use of exploratory analyses within the National Institute for Health and Care Excellence single technology appraisal process: an evaluation and qualitative analysis. *Health Technol Assess* 2016;20(26):1-48.

Appendix 1: Further critique of searches in the company submission

Clinical effectiveness

- The ERG noted that three indexing terms in line #4 of the revised clinical effectiveness strategies[13] were unnecessarily exploded, where no narrower terms existed.
- Terms used to search the conference proceedings and trial registers were overly restrictive, focussing on 'INHL' or 'indolent NHL'. Alternative terms for 'follicular lymphoma' were not used. This may have affected how sensitive these supplementary searches were, and it is possible potentially relevant references may not have been retrieved or screened.

Cost effectiveness

- The ERG noted that three indexing terms in line #4 of the revised cost effectiveness strategies[13] were unnecessarily exploded, where no narrower terms existed.
- Terms used to search the conference proceedings were overly restrictive, focussing on 'INHL' or 'indolent NHL'. Alternative terms for 'follicular lymphoma' were not used. This may have affected how sensitive these supplementary searches were, and it is possible potentially relevant references may not have been retrieved or screened.

Measurement and valuation of health effects

- The ERG noted that three indexing terms in lines #4 and #11 of these revised strategies[13] were unnecessarily exploded, where no narrower terms existed.
- Terms used to search the conference proceedings were overly restrictive, focussing on 'INHL' or 'indolent NHL'. Alternative terms for 'follicular lymphoma' were not used. This may have affected how sensitive these supplementary searches were, and it is possible potentially relevant references may not have been retrieved or screened.

Resource use

- Table 16, Line #6 (Embase strategy): The truncation symbol has been missed from a term in this line of search strategy:

6	(untreat or first?line or na?ve or "not treated" or not?treated or "not exposed" or unexposed or "new?diagnos*" or "de?novo" or "newly diagnosed" or primary or initial or early or "never?treated" or frontline or front-line or "front line" or ("without prior" and (regimen or therap* or treatment*))).ti,ab.	3907174

National Institute for Health and Care Excellence
Centre for Health Technology Evaluation

Pro-forma Response

ERG report

Obinutuzumab for untreated advanced follicular lymphoma [ID1020]

You are asked to check the ERG report from Kleijnen Systematic Reviews to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies you must inform NICE by **6pm on 28 July** 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.

Issue 1 Final licence wording

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Clarification of final licence wording at CHMP opinion.	Include the final wording at CHMP opinion in section 1.1, paragraph 2 and The wording is: "Gazyvaro in combination with chemotherapy, followed by Gazyvaro maintenance therapy in patients achieving a response, is indicated for the treatment of patients with previously untreated advanced follicular lymphoma." See: http://www.ema.europa.eu/docs/en_GB/document_library/Summary_of_opinion/human/002799/WC500231836.pdf The same should be done on p. 27, section 3.1, paragraph 2. In section 1.1, Paragraph 3, It should be noted that the final licence indication is in line with the scope. The sentence could be changed to: "The main trial in the submission, GALLIUM, excludes patients with follicular lymphoma (FL) grade 3b."	We appreciate that at the time of the submission, the final wording was not available. Including the final licence wording in the report improves clarity and provides future consistency. The latest draft SmPC at CHMP opinion is included for reference.	The final licence wording has been added.

Issue 2 PFS HR used in CS base case

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
p. 45 ERG statement "[…] instead of PFS=0.66 (0.51 to 0.85) (INV, Jan 2016)) used in the company's base-case." is incorrect.	The statement should read "[…] instead of analyses based on PFS-INV (September 2016) with a HR =0.68 (0.54 to 0.87) used in the company's base-case."	Clinical inputs from GALLIUM in the model were based on the September 2016 cut-off date thoughout the model and therefore our submission base case corresponds to the PFS-INV HR of 0.68 (0.54 to 0.87).	Data from September 2016 cut-off were used in the economic model. The PFS HR is not directly used in the model. We have adjusted the text as suggested by the company.

Issue 3 PFS HR used in ERG base case

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
p. 15 – It should be mentioned that PFS-INV was the primary endpoint and PFS-IRC a secondary endpoint.	The statement could be amended to "The *primary endpoint of* investigator (INV) assessed PFS data (PFS-INV) was used in the company's base-case analysis. The ERG considers that *the secondary endpoint of* independent review committee (IRC) assessed PFS (PFS-IRC) data should have been used for the company's base-case analysis because the GALLIUM trial was open-label […]"	For clarity, it should be stated that the CS ERGs preferred base case was based on the secondary endpoint of PFS-IRC.	Not a factual error. The endpoints of the GALLIUM trial were already discussed in the clinical effectiveness section.

Issue 4 AIC Confidentiality

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Key data from the updated analysis (clinical cut of date 10 September 2016) have been published and the AIC marking from the document can be removed.	The AIC marking from the following data presented in the document can be removed: - Date of data cut:" September 2016" (pages 12, 13, 37, 43-45, 55, 68, 74, 75, 77, 79, 114 & 118) - PFS INV HR 0.68 (table 4.6 p 44)	Data from 10Sept2016 now partially published at EHA 2017.	This has been updated.

Issue 5 Error correction in Model

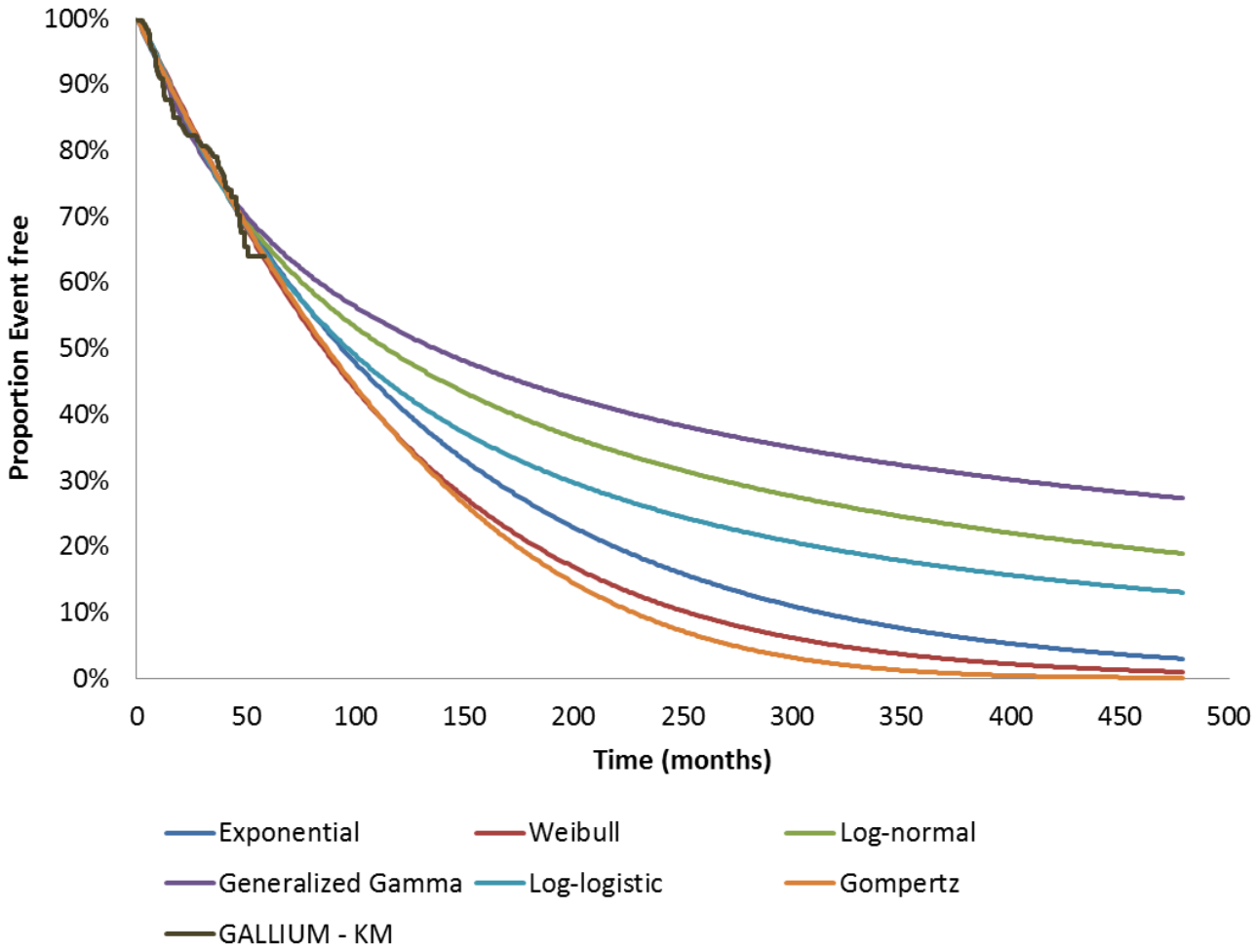
Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
ERG Model amended in 'Sheet Administration Costs' Cells G27, G35 & G36	Reverse to original calculation.	The ERG amended the transport costs to a lower value based on a cost reduction factor (Cell G19 in Sheet 'Administration Costs') used	Based on the additional explanation provided in the FEC document, the ERG agrees that the amended cells

		to adjust the maintenance costs for lower sub-cut administration costs. The patient transport costs do not need to be adjusted separately as this was already factored into the overall administration cost reduction. Moreover, MabThera Sub-cut is only funded for maintenance and therefore any reduction in administration costs for the proportion of patients receiving MabThera SC only applies in the maintenance setting. However, the effect of this change on any of the reported results is very minor.	(G27, G35 and G36) in the 'Administration Costs' sheet can be reversed to the original calculation. The ERG confirms that the effect of this change is very minor (less than 0.5% in ICER for all scenarios).

Issue 6 PFS KM incorrectly displayed

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
p. 78 Figure 5.9 incorrectly displays the KM curve for PFS-INV and not PFS-IRC. The extrapolation functions are correctly displayed (based on PFS-IRC).	Amend the figure with correct KM curve – see figure below. Remove the statement "In fact, they all seem to overestimate PFS" on p. 78.	The incorrect KM curve was displayed due to the fact the figure in the model in tab 'Tables Report' was not dynamically linked to the selected KM curve (INV or IRC) but always displayed PFS-INV data (only the figure in the 'Model Inputs' tab is linked dynamically. This can easily be amended; the correct graph is shown in the	The ERG made these statements based on the electronic model provided by the company attached to its response to the clarification letter. In the electronic model, all the figures were linked to the KM-INV data because KM-IRC curve was not available. The following amendments were done: On page 78: "*Nevertheless, from the company's electronic model, it was possible to plot the model's PFS-IRC*

		figure below.	*extrapolated curves and compare them with the KM-IRC curve for the R-chemo+R arm. This can be seen in Figure 5.9 below. Visual inspection of Figure 5.9 shows that all the parametric curves, except the generalized Gamma and the lognormal distributions, present a similar fit to the GALLIUM data. In fact, they all seem to overestimate PFS*" is changed to: "*Also, from the company's electronic model, it was not possible to plot the model's PFS-IRC extrapolated curves and compare them with the KM-IRC curves. However, the comparison for the R-chemo+R arm was done according to the KM-IRC curve provided by the company. This can be seen in Figure 5.9 below. Visual inspection of Figure 5.9 shows that all the parametric curves for the R-chemo+R arm present a similar fit to the GALLIUM data.*" We have also changed Figure 5.9 with the below figure provided by the company.


100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Proportion Event free
0
50
100
150
200
250
300
350
400
450
500
Time (months)
Exponential
Weibull
Log-normal
Generalized Gamma
Log-logistic
Gompertz
GALLIUM - KM