in collaboration with:

Bortezomib for previously untreated mantle cell lymphoma

Produced by Kleijnen Systematic Reviews Ltd (UK) in collaboration with Erasmus University Rotterdam and Maastricht University (The Netherlands)

Authors
Robert Wolff, Reviews Manager, KSR Ltd
Bram Ramaekers, Health Economist, Maastricht UMC
Annemieke van Dongen-Leunis, Health Economist, EUR
Shona Lang, Systematic Reviewer, KSR Ltd
Marianne Luyendijk, Health Economist, EUR
Remziye Zaim, Health Economist, EUR
Kate Misso, Information Specialist Manager, KSR Ltd
Gill Worthy, Statistician, KSR Ltd
Nigel Armstrong, Health Economist, KSR Ltd
Maiwenn Al, Health Economist, EUR
Johan L. Severens, Professor of Evaluation in Health Care, EUR
Jos Kleijnen, Director, KSR Ltd, Professor of Systematic Reviews ir

Correspondence to
Robert Wolff, Kleijnen Systematic Reviews
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, UK
YO19 6FD

Date completed 01/09/2015

Source of funding This report was commissioned by the NIHR HTA Programme as

project number 14/148/02.

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:
Wolff R, Ramaekers B, van Dongen-Leunis A, Lang S, Luyendijk M, Zaim R, Misso K, Worthy G, Armstrong N, Al M, Severens JL, Kleijnen J. Bortezomib for previously untreated mantle cell lymphoma: a Single Technology Appraisal. York: Kleijnen Systematic Reviews Ltd, 2015.

Contributions of authors
Robert Wolff 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. Bram Ramaekers and Maiwenn Al acted as health economic project leads, critiqued the company's economic evaluation and contributed to the writing of the report. Annemieke van Dongen-Leunis, Marianne Luyendijk, Remziye Zaim and Nigel Armstrong acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Shona Lang acted as systematic reviewer, critiqued the clinical effectiveness methods and evidence 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. Gill Worthy acted as statistician, critiqued the analyses in the company's submission and contributed to the writing of the report. Johan Severens provided senior advice and support to the cost-effectiveness section and contributed to the writing of the report. Jos Kleijnen contributed to the writing of the report and supervised the project.

TABLE OF CONTENTS

TABLE OF TABLES **6**
TABLE OF FIGURES **9**
1. Definition of terms and list of abbreviations **10**
2. Executive summary **13**
2.1 Critique of the decision problem in the company's submission 13
2.2 Summary of clinical effectiveness evidence submitted by the company 13
2.3 Summary of the ERG's critique of clinical effectiveness evidence submitted 15
2.4 Summary of cost-effectiveness submitted evidence by the company 15
2.5 Summary of the ERG's critique of cost-effectiveness evidence submitted 17
2.6 ERG commentary on the robustness of evidence submitted by the company 18
2.6.1 Strengths 18
2.6.2 Weaknesses and areas of uncertainty 18
2.7 Summary of exploratory and sensitivity analyses undertaken by the ERG 19
3. Background **20**
3.1 Critique of company's description of underlying health problem. 20
3.1.1 Mantle cell lymphoma 20
3.1.2 Prevalence of MCL 20
3.1.3 Disease burden and mortality 21
3.2 Critique of company's overview of current service provision 21
3.2.1 Marketing authorisation and health technology assessment 21
3.2.2 Current service provision and technology under assessment 22
4. Definition of the decision problem **24**
4.1 Population 25
4.2 Intervention 25
4.3 Comparators 26
4.4 Outcomes 26
4.5 Other relevant factors 26
5. Assessment of clinical effectiveness **27**
5.1 Critique of the methods of review(s) 27
5.1.1 Searches 27
5.1.2 Inclusion criteria 28
5.1.3 Critique of data extraction 29
5.1.4 Quality assessment 31
5.1.5 Evidence synthesis 36

5.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these) 36
5.2.1 Study characteristics of LYM-3002 36
5.2.2 Results 44
5.2.3 Included non-RCTs 58
5.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison 58
5.4 Critique of the indirect comparison and/or multiple treatment comparison 60
5.5 Additional work on clinical effectiveness undertaken by the ERG 62
5.6 Conclusions of the clinical effectiveness section 62
6. Assessment of cost-effectiveness 65
6.1 ERG comment on company's review of cost-effectiveness evidence 65
6.1.1 Searches for cost effectiveness review 65
6.1.2 Inclusion/exclusion criteria used in the study selection 66
6.1.3 Included/excluded studies in the cost-effectiveness review 66
6.1.4 Conclusions of the cost-effectiveness review 66
6.2 Summary and critique of company's submitted economic evaluation by the ERG 66
6.2.1 NICE reference case checklist (TABLE ONLY) 69
6.2.2 Model structure 70
6.2.3 Population 72
6.2.4 Interventions and comparators 72
6.2.5 Perspective, time horizon and discounting 74
6.2.6 Treatment effectiveness and extrapolation 74
6.2.7 Adverse events 85
6.2.8 Health related quality of life 87
6.2.9 Resources and costs 92
6.2.10 Cost-effectiveness results 106
6.2.11 Sensitivity analyses 110
6.2.12 Subgroup analysis 113
6.2.13 Model validation and face validity check 113
6.3 Exploratory and sensitivity analyses undertaken by the ERG 114
6.3.1 New ERG base case analysis 114
6.3.2 Exploratory scenario analyses undertaken by the ERG 117
6.4 Conclusions of the cost-effectiveness section 123

7. Impact on the ICER of additional clinical and economic analyses undertaken by the ERG 125
8. Conclusions 128
8.1 Statement of principle findings 128
8.2 Strengths and limitations of the assessment 128
8.3 Suggested research priorities 129
9. References 131
Appendix 1: ERG search strategies 138
Appendix 2: Summary list of cost-effectiveness evaluations 140
Appendix 3: Details of model changes implemented by the ERG 145

TABLE OF TABLES

Table 2.1: Summary of interim analysis results for LYM-3002: direct analysis of VR- CAP versus R-CHOP 13

Table 2.2: Indirect analysis of VR-CAP versus R-FC or VR-CAP versus R-bendamustine .. 14

Table 4.1: Statement of the decision problem (as presented by the company) 24

Table 5.1: Eligibility criteria applied to systematic review search results 28

Table 5.2: List of relevant RCTs for VR-CAP 30

Table 5.3: List of relevant RCTs for comparators to VR-CAP included in the final appraisal scope 30

Table 5.4: Quality assessment of the LYM-3002 trial 31

Table 5.5: Quality assessment of the European MCL Elderly trial 33

Table 5.6: Quality assessment of the StiL NHL1 trial 34

Table 5.7: Summary of the LYM-3002 trial methodology (as presented by the company) 37

Table 5.8: LYM-3002 – list of patient enrolment by region in country (as presented by the company) 39

Table 5.9: LYM-3002 – summary of statistical analyses 41

Table 5.10: LYM-3002 – overall survival (ITT set) 45

Table 5.11: LYM-3002 – overall survival (ITT set) – Western European and North American participants 46

Table 5.12: LYM-3002 – progression-free survival based on IRC assessment, ITT analysis set 47

Table 5.13: LYM-3002 – best overall response, response evaluable analysis set 48

Table 5.14: LYM-3002 – best overall response – Western European and North American participants 48

Table 5.15: LYM-3002 – time to response and duration of response, response evaluable analysis set 49

Table 5.16: LYM-3002 – Time to disease progression 50

Table 5.17: LYM-3002 – summary to disease progression, ITT set – Western European and North American participants 51

Table 5.18: LYM-3002 – safety and tolerability, safety analysis set 52

Table 5.19: LYM-3002 – overview of treatment-emergent AE by region, safety analysis set 53

Table 5.20: LYM-3002 – common AE, safety analysis set 54

Table 5.21: Risk ratios of adverse events 57

Table 5.22: Brief summary of similarity of trials included in network meta-analysis 59

Table 5.23: Bucher indirect comparison analyses of PFS and OS 61

Table 5.24: Bucher indirect comparison analyses of ORR and CR 62

Table 5.25: Summary of interim analysis results for LYM-3002: direct analysis of VR- CAP versus R-CHOP 63

Table 5.26: Indirect analysis of VR-CAP versus R-FC or VR-CAP versus R-bendamustine 63
Table 6. 1: Inclusion criteria for cost-effectiveness studies....66
Table 6.2: Summary of the company's economic evaluation (with signposts to CS)....67
Table 6.3: NICE reference case checklist....69
Table 6.4: Transition probabilities in the Markov model....74
Table 6.5: Definition of end-points....75
Table 6.6: Scenario analyses regarding the effect of maintenance treatment....82
Table 6.7: AEs weekly probabilities and duration....86
Table 6.8: Summary of QoL values used for CEA....89
Table 6.9: Disutilities for adverse events....90
Table 6.10: Number of vials per administration, number of administrations per week and costs per administration....93
Table 6.11: Dosing regimens, drug acquisition costs and administration costs....95
Table 6.12: Percentage of patients receiving first-line treatment per treatment cycle....97
Table 6.13: Difference between average dose received and per protocol dose per cycle in the LYM-3002 trial....98
Table 6.14: Distribution of second-line treatment by first-line treatment....99
Table 6.15: Medical resource use by health state....99
Table 6.16: Percentage of patients receiving concomitant medication, the dosing and weekly costs of concomitant medication....101
Table 6.17: Number of transfusions....102
Table 6.18: Concomitant treatment and transfusion costs per treatment cycle....102
Table 6.19: AE costs per week....103
Table 6.20: Concomitant treatment costs per treatment cycle (adjusted by ERG)....105
Table 6.21: Adverse event costs calculated by the ERG....105
Table 6.22: Unit prices reported in the CS that could not be reproduced by the ERG....106
Table 6.23: Disaggregated costs (discounted) per cost category and per health state....107
Table 6.24: Disaggregated QALYs and Lys (discounted) per health state....108
Table 6.25: Cost-effectiveness results....108
Table 6.26: Full incremental analysis of cost-effectiveness results....109
Table 6.27: Scenario analysis while incorporating indirect comparison for PFS and OS....112
Table 6.28: New ERG base case – pairwise comparison to VR-CAP....115
Table 6.29: New ERG base case – full incremental comparison....116
Table 6.30: Results exploratory analyses ERG; pairwise comparison VR-CAP and R-CHOP....120
Table 6.31: Results exploratory analyses ERG; pairwise comparison VR-CAP and R-FC..121

Table 6.32: Results exploratory analyses ERG; pairwise comparison VR-CAP and R-bendamustine 122

Table 7.1: Revised base case cost-effectiveness analysis, incorporating corrections and amendments identified by the ERG: VR-CAP versus R-CHOP 125

Table 7.2: Revised base case cost-effectiveness analysis, incorporating corrections and amendments identified by the ERG: VR-CAP versus R-FC 126

Table 7.3: Revised base case cost-effectiveness analysis, incorporating corrections and amendments identified by the ERG: VR-CAP versus R-bendamustine 127

TABLE OF FIGURES

Figure 5.1: LYM-3002 – study flow diagram....44
Figure 6.1: Model diagram....71
Figure 6.2: Log cumulative hazard plots for PFS....76
Figure 6.3: Fit of parametric models to PFS....77
Figure 6.4: Fit of parametric curves to OS....78
Figure 6.5: Fit of parametric models to TFI....80
Figure 6.6: Regression analysis OS from start of subsequent treatment....82
Figure 6.7: Cost-effectiveness acceptability curves....109
Figure 6.8: Tornado diagram displaying the ICER sensitivity of the 10 most influential model inputs....111
Figure 6.9: Cost-effectiveness plane for all treatment options (QALYs)....116

1. Definition of terms and list of abbreviations

aaPI	Age-adjusted International Prognostic Index
AE	Adverse event
AiC	Academic in confidence
AIC	Akaike information criterion
ALT	Alanine transaminase
ANC	Absolute neutrophils count
AS	Age-standardised
ASCO	American Society of Clinical Oncology
ASH	American Society of Hematology
AST	Aspartate transaminase
BFI	Brief Fatigue Inventory
BIC	Bayesian Information Criterion
BM	Bone marrow
BSH	British Society for Haematology
CADTH	Canadian Agency for Drugs and Technologies in Health
CDF	Cancer Drugs Fund
CEAC	Cost-effectiveness acceptability curve
CI	Confidence interval
CiC	Commercial in confidence
CR	Complete response
CRD	Centre for Reviews and Dissemination
Cru	Complete response unconfirmed
CS	Company's submission
CSR	Clinical study report
CT	Computed tomography
DNA	Deoxyribonucleic acid
DoR	Duration of response
ECOG	Eastern Cooperative Oncology Group
EHA	European Hematology Association
EMA	European Medicines Agency
eMIT	Electronic market information tool
EORTC-QLQ-C30	European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire
EQ-5D	EuroQol-5 Dimension Questionnaire
ESMO	European Society for Medical Oncology
FC	Fludarabine and cyclophosphamide
FDA	U.S. Food and Drug Administration
G-CSF	Granulocyte-colony stimulating agent
GLSG	Grade Lymphoma Study Group
HRQL	Health-related quality of life
HRQoL	Health-related quality of life
HR	Hazard ratio

HSCT	Haematopoietic stem cell transplantation
ICER	Incremental cost-effectiveness ratio
IDMC	Independent data monitoring committee
INV	Investigator
IPI	International Prognostic Index
IRC	Independent review committee
ITT	Intention-to treat
IV	Intravenous
IWRC	International Workshop Response Criteria
KM	Kaplan-Meier
LDH	Lactate dehydrogenase
LY	Life year
MCL	Mantle cell lymphoma
MIMS	Monthly index of medical specialities
MIPI	MCL International Prognostic Index
N/A	Not applicable
NE	Not estimable
NHL	Non-Hodgkin's lymphoma
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NCRI	National Cancer Research Institute
OR	Odds ratio
ORR	Overall response rate
OS	Overall survival
p_prog	Probability progression 2nd line
PAS	Patient access scheme
PD	Progressive disease
PFS	Progression-free survival
PR	Partial response
PrePS	Pre-progression survival
PRO	Patient reported outcome
PS	Performance status
PSS	Personal Social Services
QALY	Quality-adjusted life year
R-benda	Rituximab with bendamustine
R-bendamustine	Rituximab with bendamustine
R-CHOP	Rituximab, cyclophosphamide, doxorubicin, vincristine and prednisolone
R-CVP	Rituximab with cyclophosphamide, vincristine and prednisone
R-cytarabine	Rituximab with cytarabine
R-FC	Rituximab, fludarabine and cyclophosphamide
RBC	Red blood cell
RCT	Randomised controlled trial
ROW	Rest of the world

SCT	Stem cell transplant
SIGN	Scottish Intercollegiate Guidelines Network
SMC	Scottish Medicines Consortium
SmPC	Summary of Product Characteristics
SR	Systematic review
STA	Single Technology Appraisal
StiL	Study Group of indolent Lymphomas
TFI	Treatment free interval
TTNT	Time to next treatment
TTP	Time to progression
TTR	Time to response
UK	United Kingdom
ULN	Upper limit of normal
VR-CAP	Bortezomib (Velcade®) in combination with rituximab, cyclo-phosphamide, doxorubicin and prednisone
WBC	White blood cell count
WTP	Willingness to pay

2. Executive summary

2.1 Critique of the decision problem in the company's submission

The decision problem addressed in the company submission was the same as the final scope issued by NICE. The company submission report presented data which were representative of the patient population, intervention, comparators and outcomes as described in the decision problem.

2.2 Summary of clinical effectiveness evidence submitted by the company

A single ongoing study (LYM-3002) was identified for VR-CAP (bortezomib (Velcade®) in combination with rituximab, cyclo-phosphamide, doxorubicin and prednisone). The study compared the efficacy and safety of VR-CAP to that of rituximab, cyclo-phosphamide, doxorubicin, vincristine and prednisolone (R-CHOP) in patients with newly diagnosed mantle cell lymphoma (MCL) who were unsuitable for haematopoietic stem cell transplantation (HSCT). It was a phase III, randomised, open label, which recruited 487 participants from 128 sites in 28 countries. Interim analysis results (December 2013) are summarised in Table 2.1. Results for endpoints defined in the final scope are presented for the overall population and for the subgroup of Western Europe and North America (n=91).

Table 2.1: Summary of interim analysis results for LYM-3002: direct analysis of VR-CAP versus R-CHOP

Outcome from final scope	Overall Population	Western Europe and North America
Overall survival	HR = 0.80, 95% CI 0.59 to 1.10[§]	HR = 1.09; 95% CI: 0.47, 2.52[^]
Progression-free survival	HR = 0.63, 95% CI 0.50 to 0.79[□]	HR = 0.77, 95% CI 0.43 to 1.38[§]
Overall response rate	OR = 1.43, 95% CI 0.75 to 2.72[§]	OR = 1.16, 95% CI 0.32 to 4.25[§]
Duration of response (median duration, days)	VR-CAP: 1110 (95% CI 813 to 1420) R-CHOP: 459 (95% CI 379 to 518)	N/A
Time to new anti-lymphoma treatment/time to progression	HR = 0.58, 95% CI 0.45 to 0.74[$]	HR = 0.63, 95% CI 0.34 to 1.16[§]
AEs of treatment[*]: **Any TEAE** **Any SAE** **TD AE** **Deaths due to AE**	 RR = 1.01, 95% CI 0.99 to 1.03[^] RR = 1.26, 95% CI 0.98 to 1.62[^] RR = 1.25, 95% CI 0.67 to 2.30[^] RR = 0.95, 95% CI 0.49 to 1.83[§]	 RR[‡] = 1.01, 95% CI 0.97 to 1.06[^] RR[‡] = 2.20, 95% CI 1.36 to 3.56[Δ] RR[‡] = 2.14, 95% CI 0.67 to 6.80[^] N/A
Health-related quality of life (EORTC-QLQ-C30)	No difference[#]	N/A

AE = adverse event; CI = confidence interval; EORTC-QLQ-C30 = European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire; HR = hazards ratio; NA = not assessed; OR = odds ratio; OS = overall survival; R-CHOP = Rituximab, cyclo-phosphamide, doxorubicin, vincristine and prednisolone; RR = risk ratio; SAE = any serious AE; TD = Treatment discontinuation due to AE; TEAE= Any treatment-emergent AE; VR-CAP, bortezomib (Velcade®) with rituximab, cyclophosphamide, doxorubicin and prednisone

[§] In favour of VR-CAP, not statistically significant

[□] In favour of VR-CAP, statistically significant

[^] In favour of R-CHOP, not statistically significant

[Δ] In favour of R-CHOP, statistically significant

[*] AE risk ratios were calculated by the ERG

Outcome from final scope	Overall Population	Western Europe and North America
ǂ Data calculated from European Union and North America subgroups combined (n=150) # None of the observed changes in scores through the end of treatment were considered statistical significant or clinically meaningful (defined as ≥10 point change)		

For the overall population, most efficacy outcomes defined in the final scope favoured VR-CAP, but analyses were only statistically significant for progression free survival, duration of response and time to new treatment. Statistically significant differences favouring VR-CAP were also observed for time to disease progression, treatment free interval, complete response, time to response, durable overall response, and durable complete overall response. For the Western Europe and North America only four analyses defined in the final scope were reported. All of these but overall survival favoured VR-CAP but were not statistically significant. The company also reported results for overall complete response which were comparable between the two treatments (odds ratio 1.02, 95% CI 0.42 to 2.48, in favour of R-CHOP). Other outcome analyses were unavailable; more comprehensive results will be available with the final analysis in 2017.

Safety outcomes for the overall population showed that AE were more frequent in the VR-CAP compared to R-CHOP but this was not statistically significant. In the European and North America subgroup significantly more serious adverse events occurred in the VR-CAP group than in R-CHOP. Treatment emergent adverse events were similar across both groups, treatment discontinuations were higher for VR-CAP but this was not statistically significant.

No statistically significant or clinically meaningful difference regarding health-related quality of life was reported.

Three trials were identified for indirect analysis of VR-CAP versus R-FC or VR-CAP versus R-bendamustine. Results are summarised in Table 2.2. Outcomes of the final scope are presented for the overall population and for the subgroup of Western Europe and North America.

Table 2.2: Indirect analysis of VR-CAP versus R-FC or VR-CAP versus R-bendamustine

Outcome from final scope	Overall Population	Western Europe and North America
Overall survival: VR-CAP vs. R-FC VR-CAP vs. R-bendamustine	 HR = 0.53, 95% CI 0.35 to 0.81 HR = 0.63, 95% CI 0.31 to 1.25	 N/A N/A
Progression-free survival: VR-CAP vs. R-FC VR-CAP vs. R-bendamustine	 N/A HR = 1.04, 95% CI 0.59 to 1.84	 N/A N/A
Overall response rate VR-CAP vs. R-FC VR-CAP vs. R-bendamustine	 OR = 2.42, 95% CI 1.09 to 5.39 OR =3.54, 95% CI 0.61 to 20.69	 N/A N/A
Duration of response	N/A	N/A
Time to new anti-lymphoma treatment/time to progression	N/A	N/A
Adverse effects of treatment*: **Any TEAE n (%)** **Any SAE n (%)**	N/A	N/A

TD AE n (%) **Deaths due to AE n (%)**		
Health-related quality of life	N/A	N/A
AE = adverse event; CI = confidence interval; HR = hazards ratio; NA = not assessed; OR = odds ratio; OS = overall survival; R-CHOP = Rituximab, cyclo-phosphamide, doxorubicin, vincristine and prednisolone; R-FC = rituximab with fludarabine and cyclophosphamide; RR = risk ratio; SAE = any serious AE; StiL = Study Group of indolent Lymphomas; TD = Treatment discontinuation due to AE; TEAE= Any treatment-emergent AE; VR-CAP, bortezomib (Velcade®) with rituximab, cyclophosphamide, doxorubicin and prednisone		

For the overall population, only a proportion of the outcomes outlined in the scope were available for indirect analysis. Overall survival and overall response rate favoured VR-CAP in comparison to R-FC and R-bendamustine but this was only statistically significant for R-FC. Progression free survival was similar for VR-CAP and R-bendamustine.

Individual safety outcomes (Grade 3 or 4) were available for VR-CAP in comparison to R-FC. More patients who received VR-CAP compared to R-FC had lymphopenia (OR=3.81, 95% CI 1.98 to 7.31), neutropenia (OR=2.48, 95% CI 1.37 to 4.50) and thromobocytponeia (OR=9.19, 95% CI 3.99 to 21.17). Similar numbers of patients had anaemia, leukopenia or fibrile neutopenia between the two groups.

2.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

The risk of bias of the only RCT, LYM-3002, assessing VR-CAP as well as the two additional studies included in the indirect comparison was judged to be high. Therefore, results should be interpreted with caution.

For LYM-3002 only interim results (December 2013) are available, the final analysis is planned for 2017. Furthermore, this trial did not include any participants from the UK and only approximately one third of the participants were from Europe or North America. Focussing on this subgroup, results for all endpoints will be underpowered.

The majority of searches in the company submission were well documented and easily reproducible; searches were carried out in line with the NICE guide to the methods of technology appraisal sections 5.2.2 and 5.2.4. In addition, relevant references have been excluded. However, as these were linked to the LYM-3002 trial, no relevant data were excluded.

2.4 Summary of cost-effectiveness submitted evidence by the company

The company developed a *de novo* cost-effectiveness model to assess the cost-effectiveness of VR-CAP for the treatment of patients with previously untreated MCL that are unsuitable for HSCT. R-CHOP was chosen as the base-case comparator as this is the current standard care in UK practice. Other treatments that are used in the UK are R-FC and R-bendamustine and these treatments were chosen as comparators for exploratory analyses.

The model was a Markov health state model and included five different health states: 'progression-free survival (PFS) from first-line treatment', 'progressed from first-line treatment', 'progression-free survival from second-line treatment', 'progressed from second-line treatment' and death. All patients start in the 'PFS from first-line treatment' health state. Subsequently, patients may experience different phases of the disease from 'progression from first-line' to 'PFS from second-line treatment' and 'progression from second-line treatment'.

The model distinguishes ‘progressed from first-line treatment’ from ‘PFS from second-line treatment’ as second-line treatment does not start immediately after disease progression. In all health states patients are at risk to die. The cycle length of the model was one week and the time horizon of the study was 20 years.

Data from the LYM-3002 trial were used to estimate transition probabilities between the health states, based on time-to-event curves for Overall Survival (OS), PFS, and the ‘Treatment Free Interval’ (TFI). One exception was made, i.e. for the transition from ‘PFS from second-line treatment’ to ‘progressed from second-line treatment’. That transition probability was derived from the average duration of ‘PFS from second-line treatment’ in LYM-3002.

Utilities for ‘PFS from first-line treatment’, ‘progressive disease from first-line treatment’ and utility decrements of adverse events (AEs) of first-line treatment were derived from EQ-5D data from the LYM-3002 trial. The LYM-3002 trial did not capture utilities of second-line treatment and therefore utilities for ‘PFS from second-line treatment’ and progressive ‘disease from second-line treatment’ were based on literature and assumptions based on clinician input. AEs of second-line treatment were not included in the model.

Categories considered for resource use and costs were: drug acquisition and administration costs, health state costs, adverse event costs and end of life costs. The average number of vials/units required per administration and cost per administration for separate drugs were calculated based on an average body surface area of patients from the US, Canada and Western European in the LYM-3002 trial. Health state costs attributed to disease management comprised of the costs related to tests, scans, medical visits, concomitant medication and transfusions. Adverse event costs were calculated based on weekly probabilities and unit costs (NHS reference costs). End of life costs were based on the average costs of care for patients with cancer in the last eight weeks of life.

The base case incremental cost-effectiveness ratio (ICER) for VR-CAP compared to R-CHOP was £20,362 per quality adjusted life year gained. ICERs versus other comparators were £18,509 versus R-FC and £13,797 versus R-bendamustine. One-way deterministic sensitivity analysis, scenario analyses and probabilistic sensitivity analysis analyses were conducted.

From the deterministic sensitivity analysis the company concluded that ICERs were most sensitive to the survival functions used to model PFS and OS, the utility for patients progressed from second-line treatment, IV administration costs and the duration of second-line treatment. The scenario analysis showed that the ICER increased when different parametric distributions for PFS were incorporated, the utility value for patients progressed from second-line treatment was increased, all health state utility values were decreased (except for the progressed from second-line treatment utility, this was kept the same), no second-line treatment was assumed, R-maintenance was included, the time horizon was changed to 10 years and baseline age of 69 was set to 75.

2.5 Summary of the ERG's critique of cost-effectiveness evidence submitted

The economic model described in the CS is considered by the ERG to meet the NICE reference case to a reasonable extent. The ERG also confirmed the company's finding that there was no existing cost-effectiveness model for bortezomib for the current indication.

In terms of comparators, the CS did not include R-bendamustine and R-FC in the base case analysis, but presented the results of these treatments as 'exploratory' only, despite the fact that these two comparators were mentioned in the scope. The company justified this decision by explaining that R-CHOP is established standard of care whereas R-bendamustine and R-FC are used only for patients unsuitable due to frailty as a result of advanced age and/or comorbidities. The ERG thinks that exclusion of R-FC and R-bendamustine in the base case is inconsistent with the scope. Therefore, the ERG included these comparators in all analyses, not only as 'exploratory comparators'. In addition, the ERG thinks that the exclusion of post induction R-maintenance treatment is debatable since at the scoping meeting clinical experts indicated that maintenance treatment is common practice in the UK.

The company model follows a logical structure with respect to the nature of the disease. However, the ERG did not agree with the company to not apply half cycle correction and to use the ITT population of the LYM-3002 trial to assess the effectiveness of VR-CAP versus R-CHOP instead of the European Union subgroup. The treatment effectiveness in the LYM-3002 trial seems lower for the European Union population and thus using this subgroup would result in more conservative estimates.

In general, the ERG has several comments on the way the PFS, OS and TFI were estimated. First, it is uncertain whether the best fitted model was selected to model the PFS for the VR-CAP arm. The log-logistic distribution was selected whereas the exponential distribution had the best statistical fit. Another concern regarding the estimation of PFS is the use of a stratified model whereas an unstratified model seemed to be more appropriate. Second, regarding the modelling of survival, the ERG questions the use of different survival curves based on progression status and the assumption that survival for patients without progression differs between treatment arms. Third, the ERG expects that TFI is overestimated by the company and possibly this overestimation is larger in VR-CAP patients since these estimates are based on all patients irrespective of treatment response.

The ERG identified a model error in the calculation of AE cycle probabilities, and disagreed with some of the other AE related input values. However, these issues had only little impact on the ICERs.

Utilities for second-line treatment were based on assumptions and literature. The ERG agrees with the assumption that the utility for 'PFS from second-line treatment' is similar to the 'PFS from first-line treatments' however does not agree with the value used for 'progressed disease from second-line treatment'. This utility is based on a study about aggressive NHL and is clearly uncertain given the small patient population on which this estimate is based. The ERG also noted that some utility values used for the secondary scenario analyses were not correctly entered into the company model.

The ERG agreed with using the mean second-line duration for drug acquisition and administration as this was seen as conservative compared to using the maximum second-line duration. In contrast, the ERG did not agree with the dose reduction applied to the drug costs for VR-CAP and R-CHOP since it is questionable whether the dose reduction observed in the LYM-3002 trial is representative of UK clinical practice. In addition, some errors in the model were found regarding concomitant medication costs and costs for pegfilgrastim for R-CHOP.

Finally, the ERG did not agree with using the end-of-life costs based on only 40 cancer patients receiving palliative care as it is questionable whether these costs are representative of the MCL population. Moreover, this cost estimate included costs for 'third sector' (i.e. not-for-profit and non-governmental) healthcare organisations.

2.6 ERG commentary on the robustness of evidence submitted by the company

2.6.1 Strengths

Overall, the CS is well presented and in line with the final scope.

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 submission and response to clarification provided sufficient details for the ERG to appraise the searches. Additional searches were carried out for conference abstracts.

For the health economic evaluation, the impact of assumptions was extensively explored in the scenario analyses. In addition, the current model contained utility values measured (using EQ-5D) alongside the LYM-3002 study for most health states.

2.6.2 Weaknesses and areas of uncertainty

The only RCT assessing bortezomib, LYM-3002, did not include any patients from the UK and only a third of the participants were from Europe or North America. Given that prevalence varies by geographic region and potential geographic differences in clinical standards (e.g. concomitant care), the question of generalisability arises.

Due to risk of bias, results should be interpreted with caution. In addition, analyses of many endpoints are underpowered, especially when focussing on the subgroup of European and North American participants.

The ERG was concerned that specific adverse events searches without the restriction of a study design filter were not conducted; this is not in line with current best practice. Searches of the Cochrane Library for sections 5.4 and 5.5 were overly restrictive and may have impaired recall. Also, the dates of clinical effectiveness searches were reported inconsistently in Appendices 3 and 15 of the company submission.

Most uncertainty in the health economic model was related to the estimation of overall survival. The ERG questions the validity of approach to separately model OS for patients with and without progression and the assumption that initial treatment only impacts OS for patients without progression.

2.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

A new base case was identified by the ERG, increasing the ICER of VR-CAP versus R-CHOP by £14,000 to £34,039 compared to the CS base case, which represents a 75% increase. The ICER of VR-CAP versus R-FC changed from £18,430 to £13,455 and the ICER of VR-CAP versus R-bendamustine increased from £13,725 to £16,762. In the full incremental analysis, where all treatments are compared together, it was shown that R-FC and R-bendamustine are dominated by R-CHOP i.e. they had higher costs and lower QALYs.

In addition, the PSA showed that the probability that VR-CAP is cost-effective is much smaller in the ERG base-case compared to the company's base-case (11% versus 49% and 39% versus 89% for a threshold of £20,000 and £30,000, respectively). Similar to the company's base case, the probabilities that R-FC or R-bendamustine are cost-effective at these thresholds are negligible.

This large difference between the company base case and the ERG base case was caused mainly by changing the distribution for PFS in the VR-CAP arm to the exponential distribution, whilst keeping the distribution for R-CHOP PFS a log-logistic distribution.

In the exploratory ERG analyses that were done using the ERG base case as starting point, the most extreme ICER was observed when the overall survival of VR-CAP was assumed to be similar to the overall survival of R-CHOP: £328,757. The next highest ICER was found when R-maintenance is included in the model, yielding an ICER of £43,779 for VR-CAP versus R-CHOP.

3. Background

3.1 Critique of company's description of underlying health problem.

This report provides a review of the evidence submitted by Janssen in support of bortezomib (trade name Velcade®) for the treatment of mantle cell lymphoma (MCL) in patients who are previously untreated and are unsuitable for haematopoietic stem cell transplantation (HSCT).

3.1.1 Mantle cell lymphoma

The final scope stated that *"Mantle cell lymphoma (MCL) is a rare and often aggressive type of non-Hodgkin's lymphoma which affects B-cells".*[1]

According to page 24 of the company's submission (CS)[2], *"MCL is a rare and often aggressive sub-type of B-cell lymphoma that represents approximately 6% of all NHL* [non-Hodgkin's lymphoma] *cases*."[CS references 10-12]

MCL is a genetic disorder which results in the deregulation of the cell cycle and DNA (Deoxyribonucleic acid) damage response pathways. The result of cellular deregulation is the abnormal proliferation of B cells.[CS reference 15] At presentation, tumours are found predominantly in the lymph nodes (75%), spleen (massive splenomegaly in 45–60% of cases), liver (hepatomegaly in 35%), Waldeyer's ring, bone marrow (>60%), blood (13–77%) and extranodal sites (especially the gastrointestinal tract).[CS reference 17] This tumour is considered one of the most aggressive lymphoid neoplasms with poor responses to conventional chemotherapy and relatively short survival (median overall survival of patients with MCL is 3–4 years). A subset of patients with a more indolent clinical course has been recognised.

ERG comment: The company submission report presents an accurate description of the disease. It is notable that the clinical evolution of patients with MCL is very variable and overall it is very aggressive but there is a subset of patients with a more indolent clinical course that may not require treatment. According to the CS[2], "*these patients do not require active treatment and enter the decision problem of interest to this submission when their disease course becomes more aggressive*."

3.1.2 Prevalence of MCL

The final scope stated that *"Approximately 10,800 people were diagnosed with non-Hodgkin's lymphoma in England in 2011, including approximately 500 with mantle cell lymphoma. Mantle cell lymphoma is more common in men than women (75% of people with mantle cell lymphoma are men), and it predominantly affects older people (the median age at presentation is 63 years). Most people with mantle cell lymphoma are diagnosed in advanced stages of the disease".*[1]

According to page 24 of the CS[2], *"NHL describes a diverse group of blood cancers characterised by the abnormal proliferation of malignant lymphocytes; this group accounts for approximately 3-4% of all cancer cases worldwide.*.[CS references 8+9] *MCL is a rare and often aggressive sub-type of B-cell lymphoma that represents approximately 6% of all NHL cases."*[CS references 10-12]

ERG comment: The ERG finds that the reported rates and percentages for non-Hodgkin's lymphoma and MCL are correct. According to Cancer Research UK[3] it was expected that 513 new cases of MCL would occur in the UK, which equates to an Age-Standardised (AS) Incidence Rates of 0.6 per 100 (estimates based on Haematological Malignancy Research Network data, 2004-2011).

3.1.3 Disease burden and mortality

According to page 25 of the CS[2], "*the general pattern of disease progression in MCL is one of relapse and remission, with each relapse becoming more difficult to treat, and the depth and durability of any subsequent remissions achieved invariably inferior to those achieved with first-line treatment.*"[CS references 7, 17, 22, 34, 38]

According to page 28 of the CS[2], *"the incidence rate of MCL in the UK is reported to be 0.9 per 100,000.*[CS reference 53] *Based on the Office for National Statistics 2015 population estimates for England and Wales (57 million).*[CS reference 54], *approximately 500 patients are estimated to be diagnosed with MCL in England and Wales each year (see section 6). Of these patients, approximately two thirds are likely to present with aggressive disease requiring treatment but for whom HSCT is unsuitable, and of those, a further two thirds are expected to be fit for CHOP-like therapy.*[CS reference 86] *An estimated 215 patients in NHS England and 12 patients in NHS Wales would be eligible for VR-CAP therapy (see section 6).*"[CS reference 6]

ERG comment: No data were provided regarding mortality rates. The ERG found the following information[4]: which states:

"The median overall survival time for MCL is shorter than that seen in other lymphomas, at approximately 3-5 years.[CS reference 7] *Some patients follow a very rapid progression, dying of the disease a few months after diagnosis, whereas others have an indolent course and survive more than 10 years. Approximately 30% of patients have complete response to current treatments. 232 deaths from MCL were registered in England and Wales during 2010-2011 (C83.1).*"[CS reference 9]

From this document the number of MCL deaths was obtained from the Office of National Statistics and found to be accurate. These data further identified that 170 deaths took place in males and 62 in women. The greatest number of male deaths (30) occurred in the age range 75-79 years and the greatest number of female deaths (13) occurred in the age range 70-74 years. The youngest age of death from MCL was found in the age range of 40-44 years (two males). 2013 statistics were also available and demonstrated that 242 deaths occurred due to MCL.[5]

3.2 Critique of company's overview of current service provision

3.2.1 Marketing authorisation and health technology assessment

According to page 19 of the CS[2], VR-CAP was approved by the European Commission for a marketing authorisation on 2 February 2015 for the indication detailed in this submission. This approval was based on the trials LYM-30 02, LYM-2034 and M34103-053. In addition the CS states that *"there are no restrictions on use, but contraindications of hypersensitivity*

to the active substance or to any listed excipients and of acute diffuse infiltrative pulmonary and pericardial disease are noted."

According to page 19 of the CS[2], *"outside of Europe, bortezomib has marketing authorisation for the treatment of adult patients with previously untreated MCL who are unsuitable for HSCT in US, Canada, Costa Rica, Egypt, El Salvador, Iran, Lebanon, Peru and Syria. (...) In the UK, an application to the Scottish Medicines Consortium (SMC) was submitted on 2 April 2015, and final advice is expected in September 2015."*

ERG comment: The ERG has no specific comments regarding marketing authorisation and health technology assessment.

3.2.2 Current service provision and technology under assessment

There is currently no published NICE guidance on the treatment of MCL. In 2012, a UK MCL guideline group published guidelines for the investigation and management of MCL[6] and this is frequently referred to in the CS.

The CS reports on page 21 that *"Bortezomib does not require any additional tests or investigations beyond those used to confirm MCL"*.[2]

On page 22 of the CS *"administration of bortezomib would utilise this existing NHS infrastructure. (...) Bortezomib requires administration on two additional days within the 21-day treatment cycle compared with vincristine, the agent bortezomib replaces in the R-CHOP regimen.*[CS reference 1] *This represents the main additional resource use to the NHS"*

The final scope stated that *"Mantle cell lymphoma has been one of the most difficult types of non-Hodgkin's lymphoma to treat. Although it often responds well to initial chemotherapy, the duration of remission is often short and the median overall survival is 3–5 years. There is no accepted standard of care for mantle cell lymphoma, and the choice of treatment depends on the overall aim of therapy, the grade of disease, age and fitness"*.[1]

According to page 25 of the CS[2], *"there are no strict criteria against which patients are assessed; rather, haematologists will assess eligibility on a patient by patient basis, taking into account factors, such as patient age; performance status and disease prognosis (assessed within the MCL International Prognostic Index [MIPI]); disease severity; co-morbidities status; and clinical risk."*[CS references 7, 20, 34, 35, 38-40]

"For patients who are not eligible for high-intensity induction therapy, that is those for whom HSCT is unsuitable, there are no licensed induction therapy regimens. UK guidelines recommend treatment with a rituximab-based chemotherapy induction regimen[CS reference 7], *and whilst there is still no unanimously accepted front-line option within these chemotherapeutic regimens, since the large scale European MCL Elderly trial demonstrated a survival benefit for R-CHOP*[CS reference 41], *this regimen has become the preferred first-line induction therapy option in UK clinics."*

ERG comment: According to the guidelines quoted by the CS[6], options for chemotherapy in MCL patients whom HSCT is unsuitable include R-CHOP, R-CVP, R-FC, R-bendamustine and R-chlorambucil. This indicates a greater number of alternative treatments than those discussed in the CS section of 'clinical pathways of care'. It is noted from the advisory board

summary report[7] that R-CHOP is the preferred first line treatment in the UK because it has the best evidence base to support its clinical benefit. Options for the alternative treatments are given below:

- *"R-CVP is the most commonly used alternative [to R-CHOP]"*. R-CVP has lower toxicity than R-CHOP and is therefore selected when patients are not considered fit enough to receive R-CHOP.[7]
- *"R-bendamustine is only available through the Cancer Drugs Fund (CDF) for patients unsuitable for standard first-line treatment."*[CS reference 42][2]
- *"R-FC is listed in treatment guidelines as a potential treatment option for newly diagnosed MCL patients, in the UK this induction regimen is rarely used in the front-line setting due to a poorer survival benefit than R-CHOP*[CS reference 41] *and concerns regarding the long-term impact of R-FC on myelosuppression which compromises the ability to deliver further treatments at relapse.*"[7]
- R-chlorambucil is only used in *"very frail, elderly patients"*.[7]

Recent reviews recognise the heterogeneity of the disease (a subset of patients have indolent clinical course) and suggest that therapies should be tailored to the genetic/cellular characteristics as well as the clinical characteristics of the disease.[8]

4. Definition of the decision problem

Table 4.1: Statement of the decision problem (as presented by the company)
(Table 1 on page 13-14 of the CS[2])

Key parameter	Final scope issued by NICE	Decision problem addressed in the submission	Rationale if different from the scope
Population	People with previously untreated mantle cell lymphoma, who are not going to have a stem cell transplant.	As defined by scope	N/A
Intervention	Bortezomib in combination with rituximab, cyclo-phosphamide, doxorubicin and prednisone (VR-CAP).	As defined by scope	N/A
Comparator(s)	• R-CHOP (rituximab, cyclo-phosphamide, doxorubicin, vincristine and prednisolone)* • bendamustine plus rituximab (with or without cytarabine) • R-FC (rituximab, fludarabine and cyclophosphamide)	As defined by scope	N/A
Outcomes	• overall survival • progression-free survival • overall response rate • duration of response/remission • time to new anti-lymphoma treatment/time to progression • adverse effects of treatment • health-related quality of life	As defined by scope	N/A
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.	A cost-effectiveness analysis expressed in terms of incremental cost per quality-adjusted life year is presented. A lifetime time horizon of 20 years is	N/A

Key parameter	Final scope issued by NICE	Decision problem addressed in the submission	Rationale if different from the scope
	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.	used in the base case analysis. Costs are considered from an NHS and Personal Social Services perspective.	
Subgroups to be considered	None reported	N/A	N/A
Special considerations, including issues related to equity or equality	None reported	N/A	N/A

HRQoL = health-related quality of life; N/A = not applicable; NICE = National Institute for Health and Care Excellence; R-CHOP = rituximab, cyclophosphamide, doxorubicin, vincristine and prednisolone; R-FC = rituximab, fludarabine and cyclophosphamide; VR-CAP = Bortezomib (Velcade®) in combination with rituximab, cyclophosphamide, doxorubicin and prednisone

* NB: Vincristine is not listed in table 1 of the CS, likely due to an error in the final scope where the definition of R-CHOP in the main table but not the background text did not include vincristine as part of the R-CHOP regimen.

4.1 Population

The patient population described in the final scope is as follows: *"People with previously untreated mantle cell lymphoma, who are not going to have a stem cell transplant."*[1] This is in line with the patient population included in the CS.[2]

ERG comment: The table above is based on the final scope issued by NICE.[1] Overall, the ERG is convinced that the population is in line with the final scope. However, it should be noted that the trial identified for the comparison of bortezomib with R-CHOP (LYM-3002[9]) include patients that might be outside the definitions used in the final scope. This point is discussed in section 5.2.1.

4.2 Intervention

The intervention described in the final scope is as follows: *"Bortezomib in combination with rituximab, cyclophosphamide, doxorubicin and prednisone (VR-CAP)."*[1]. This is in line with the intervention included in the CS.[2]

ERG comment: The intervention in the CS matches the intervention described in the final scope.

4.3 Comparators

The comparators described in the final scope were as follows: *"R-CHOP (rituximab, cyclophosphamide, doxorubicin and prednisolone)*[1] *or bendamustine plus rituximab (with or without cytarabine) or R-FC (rituximab, fludarabine and cyclophosphamide)."*[1] This is in line with the comparators included in the CS.[2]

ERG comment: The table above is based on the final scope issued by NICE. Overall, the ERG is convinced that the comparators are in line with the final scope. Indirect comparisons were carried out by the company to other alternative therapies (see section 4.10 of the CS).[2] Other potential comparators are R-CVP and R-chlorambucil, neither was considered since were not considered or included in the final scope (discussed in section 5.4 of the ERG critique).

4.4 Outcomes

The outcomes described in the final scope were as follows: *"overall survival, progression-free survival, overall response rate, duration of response/remission, time to new anti-lymphoma treatment/time to progression, adverse effects of treatment, health-related quality of life".*[1]. This is in line with the outcomes included in the CS.[2]

ERG comment: The ERG is convinced that the outcomes are in line with the final scope. The direct analysis (VR-CAP versus R-CHOP) provided by the LYM-3002 trial reports data for all outcomes.

4.5 Other relevant factors

ERG comment: No patient access scheme (PAS) was included in the CS. End of life criteria are not relevant for this project.

[1] R-CHOP also includes vincristine as detailed in the note of table 4.1

5. Assessment of clinical effectiveness

5.1 Critique of the methods of review(s)

5.1.1 Searches

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.[10] The submission was checked against the Single Technology Appraisal (STA) company evidence submission template and user guide.[11, 12] The ERG has presented only the major limitations of each search strategy in the main report. Further criticisms of each search strategy can be found in Appendix 1.

Clinical effectiveness/identification of studies

The company submission stated that a comprehensive review of the published literature was carried out. Three sets of searches were conducted to identify randomised controlled trials (RCT), systematic reviews (SR)/meta-analyses and non-randomised studies (non-RCT). Strategies were reported for MEDLINE, MEDLINE In-Process, Embase and the Cochrane Library. These meet the requirements specified in current best practice guidance as detailed in the NICE guide to the methods of technology appraisal.[13] No date or language limits were applied. The company performed additional searches for the following conference proceeding websites: the American Society of Hematology (ASH), the British Society for Haematology (BSH), the European Hematology Association (EHA), the American Society of Clinical Oncology (ASCO) and the European Society for Medical Oncology (ESMO). The conference proceeding searches were also well reported and clearly documented.

The database hosts for each database, date span and search dates were listed. Update search dates and numbers retrieved were also provided. The searches were on the whole well-reported and reproducible. The database searches were clearly structured and used combinations of index terms and free text. The study design limits applied to identify RCTs, SRs and non-RCTs were applied appropriately and appeared to be those developed by the Scottish Intercollegiate Guidelines Network (SIGN).[14] The clinical effectiveness searching was also reported in Appendix 15,[15] which presented further work looking at surrogate outcomes. There were discrepancies in the search dates reported in Appendices 3 and 15,[15] which were not clarified in the clarification response.[16]

Additional work on surrogate outcomes was reported in Appendix 15.[15]

Indirect and mixed treatment comparisons

The clinical effectiveness searches reported in section 4.1 and Appendix 3 of the CS[2] were used to inform the indirect and mixed treatment comparisons. As these searches did not include an intervention facet, the ERG considered the searches adequately sensitive and fit for purpose.

Non-randomised and non-controlled evidence

The non-RCT searches were reported within section 4.1 of the company submission and are discussed above with the clinical effectiveness searches.

Adverse events

No specific AE searches were performed. When the ERG queried this omission, the company responded that the clinical effectiveness searches were considered broad enough to capture studies of any design. Guidance by the Centre for Reviews and Dissemination (CRD) 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.[17] Despite three separate searches combining filters to limit to systematic reviews, RCTS and non-randomised studies, the ERG considered that it was possible that some relevant evidence may not have been identified as a consequence of the study design limits. Unfortunately the ERG was unable to undertake independent AE searches and complete a full systematic review of AEs within the STA timeline, as this would be outside of the ERG remit.

Summary of searching

The majority of searches in the CS were well documented and easily reproducible; searches were carried out in line with the NICE guide to the methods of technology appraisal sections 5.2.2 and 5.2.4.[13] Separate adverse events searches were not conducted.

5.1.2 Inclusion criteria

According to page 32 of the CS *"clinical trials of any design were included in the final evidence base of relevant studies as long as they investigated the clinical efficacy and/or safety of interventions named in the decision problem for the first-line induction therapy of adult patients with MCL and reported outcomes prior to consolidation with HSCT if this was subsequently conducted"*.[2]

Table 5.1 shows eligibility criteria applied to systematic review search results.

Table 5.1: Eligibility criteria applied to systematic review search results
(Table 7 of the CS[2])

	Inclusion criteria	Exclusion criteria
Study population	Adult patients MCL (active disease) No prior therapy for MCL	Paediatric patients Non-MCL patients Relapsed/refractory MCL patients Patients in remission
Interventions	VR-CAP R-CHOP R-FC R-bendamustine	Listed interventions with HSCT consolidation[a] Any other active therapy
Comparators	Any active therapy Placebo No treatment	None
Outcomes	DoR HRQL OS PFS Response rates Safety/tolerability TTP TTNT	None

	Inclusion criteria	Exclusion criteria
Study design	RCTs Non-RCTs Non-controlled trials (single-arm) Prospective observational studies Retrospective observational studies	In vitro studies Case report Case studies/series Letter Commentary Editorial
Key: DoR, duration of response; HRQL, health-related quality of life; HSCT, haematopoietic stem cell transplantation; MCL, mantle cell lymphoma; OS, overall survival; PFS, progression-free survival; R-bendamustine, rituximab with bendamustine; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisone; RCT, randomised controlled trial; R-FC, rituximab with fludarabine and cyclophosphamide; TTNT, time to next treatment; TTP, time to progression; VR-CAP, bortezomib (Velcade®) with rituximab, cyclophosphamide, doxorubicin and prednisone. **Notes:** [a], unless outcomes prior to consolidation with HSCT are reported.		

ERG comment:

- The population is defined as 'Adult patients with active MCL and no prior therapy for MCL'. This was not in line with the scope which additionally specified that patients should not be going to have a stem cell transplant. The CS (page 32) noted that '*no restriction regarding suitability for HSCT was applied as this is assessed on a patient by patient basis in clinical practice (see section 3.2) and not a common criterion applied in clinical trial protocols.'* As discussed in section 5.2.1, the included LYM-3002 trial enrolled "*patients ineligible or not considered for HSCT"*.[2]
- The intervention is defined to be *"VR-CAP, R-CHOP,* R-bendamustine, R-FC*"* which does not reflect the final scope which defines the intervention as *"VR-CAP"*. However, the inclusion criteria for the systematic review are appropriate with those required for a network analysis whereby all interventions and comparators are considered 'interventions' and any active therapy, placebo and no treatment are considered the 'comparators'. The review considered network analysis dependent on the available evidence. Therefore ERG considers the interventions and comparators as appropriate.
- Relevant study designs included *"RCTs, Non-RCTs, Non-controlled trials (single-arm), Prospective observational studies, Retrospective observational studies"*. This was not specified in the final scope but seems justified.
- The specified outcomes of the inclusion criteria match those of the scope.

5.1.3 Critique of data extraction

Studies identified from the searches were screened initially using a single reviewer to screen on the basis of population (MCL) and study design. This was justified on the basis that '*a number of clearly irrelevant citations were identified due to the broad information retrieval methods*'.[2] This is not a recommended method.[17, 18] This initial screen was performed on 4,762 articles and excluded 3,757 references. The remaining 1,005 articles were appropriately screened by two reviewers and 414 full text articles were also screened by two reviewers. A third reviewer was used in the event of disagreements.

According to Figure 2 of the CS, 18 studies (reported in 25 articles) met the inclusion criteria, whilst 391 full text articles (including two additional records identified through other sources) were excluded. Reasons for exclusion were all appropriate to the review. Two sources of

unpublished data were also included in the final evidence base: a clinical study report by Janssen[CS reference 60] and subgroup analyses the Study Group of indolent Lymphomas (StiL) provided by the primary author on request[CS reference 61].

Seven included studies were for randomised controlled trails (summarised in Tables 5.2 and 5.3) and 11 non-randomised or non-controlled trials (summarised in Appendix 10 of the CS).

No detailed methods were presented for extraction procedure.

Table 5.2: List of relevant RCTs for VR-CAP
(Table 8 of CS[2])

Trial Number	Population	Intervention	Comparator	Primary Study Reference	Secondary Study References
LYM-3002	Adult patients with previously untreated MCL unsuitable for HSCT.	VR-CAP	R-CHOP	Robak et al. 2015[9]	Drach et al. 2014[19] Robak et al. 2014[20] LYM-3002 CSR[21]
Key: CSR, clinical study report; HSCT, haematopoietic stem cell transplantation; MCL, mantle cell lymphoma; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisone; RCT, randomised controlled trial; VR-CAP, bortezomib (Velcade®) with rituximab, cyclophosphamide, doxorubicin and prednisone.					

Table 5.3: List of relevant RCTs for comparators to VR-CAP included in the final appraisal scope
(Table 23 of CS)

Trial Name	Population	Intervention	Comparator	Primary Study Reference	Secondary Study References
LY05	Previously untreated MCL	R-FC	FC	Eve et al. 2009[22]	Rule et al. 2005[23]
BRIGHT	Previously untreated indolent NHL or MCL	R-bendamustine	R-CHOP/ R-CVP	Flinn et al. 2014[24]	-
European MCL Elderly trial	Aged over 66 years with newly diagnosed MCL; ineligible for HSCT	R-FC	R-CHOP	Kluin-Nelemans et al. 2012[25]	-
GLSG	Previously untreated follicular lymphoma, MCL or lympho-	R-CHOP	CHOP	Lenz et al. 2005[26]	Hoster et al. 2008[27]

Trial Name	Population	Intervention	Comparator	Primary Study Reference	Secondary Study References
	plasma-cytic lymphoma; eligible for HSCT				
NCRI III	Newly diagnosed MCL	R-FC	FC	Rule et al. 2011[28]	-
StiL NHL	Newly diagnosed indolent lymphoma or MCL	R-bendamustine	R-CHOP	Rummel et al. 2013[29]	Rummel et al. 2014[30] Rummel et al. 2014[31]

Key: FC, fludarabine and cyclophosphamide; GLSG, Grade Lymphoma Study Group; HSCT, haematopoietic stem cell transplantation; MCL, mantle cell lymphoma; NCRI, National Cancer Research Institute; NHL, non-Hodgkin's lymphoma; R-bendamustine, rituximab with bendamustine; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisone; R-CVP, rituximab with cyclophosphamide, vincristine and prednisone; R-FC, rituximab with fludarabine and cyclophosphamide; RCT, randomised controlled trial; StiL, Study Group of indolent Lymphomas; VR-CAP, bortezomib (Velcade®) with rituximab, cyclophosphamide, doxorubicin and prednisone.

ERG comment: It is not a recommended method to perform any screening with a single reviewer; therefore it is a weakness in the methodology of the review.[17, 18] Appendix 4 of the CS lists the citations excluded at tertiary screening stages.[2] In this list, the ERG identified three references linked to LYM-3002.[32-34] It is unclear why these references were excluded and not included as further secondary study references in Table 8 of the CS (cf. Table 5.3 above). While these three references do not add anything to the four references already identified in the CS, it could be seen as an indication that the process of screening studies for inclusion might has been flawed, i.e. that further potentially relevant studies might have been excluded. Methods for data extraction were unclear.

5.1.4 Quality assessment

Five studies were included in the network analysis for clinical effectiveness, but only three (LYM-3002, European MCL Elderly, StiL NHL1) provided data appropriate to interventions or comparators of the scope. The decision to include these studies in the analysis is discussed later in Section 5.4.

The quality assessments of LYM-3002, European MCL Elderly trial and StiL NHL1 were presented in Tables 22-24 of Appendix 7 of the CS.

Table 5.4: Quality assessment of the LYM-3002 trial
(Based on Appendix 7, Table 22 of the CS[2])

Study question	How is the question addressed in the study?	Risk of bias	ERG comment
Was randomisation carried out appropriately?	Randomisation performed by permutated blocks method and stratified	Low	Agree. Based on a computer-generated randomisation.

Study question	How is the question addressed in the study?	Risk of bias	ERG comment
	according to IPI and stage of disease at diagnosis.		
Was the concealment of treatment allocation adequate?	Randomisations performed at a central data centre.	Low	Agree
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Patient demographics well balanced; prognostic factors reported including age, disease stage, LDH and BM involvement similarly distributed. Slight imbalance between the distribution of subjects across ECOG 0-1 but disease staging distribution similar.	Low	Moderate. There is an imbalance in ECOG as reported by the CS. In addition, there is a greater proportion of Asian patients receiving VR-CAP (36.2%) than R-CHOP (27.9%)
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)?	Open-label study with care providers and participants not blinded due to different dose schedules for bortezomib and vincristine. Outcome assessors were blind to treatment allocation with efficacy assessed by central review of radiology. Sponsor and investigational sites blinded to all analyses reviewed by the IDMC.	Low	This should be 'High' because the trial is open label
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Discontinuation rate comparable across treatment groups. Evaluable patient numbers comparable across treatment groups.	Low	Agree
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No.	Low	Agree

Study question	How is the question addressed in the study?	Risk of bias	ERG comment
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Primary efficacy analysis was performed according to the intention-to-treat principle with standard censoring methods used to account for missing data.	Low	Agree
Was statistical powering such to detect a significant difference between treatment groups?	Yes.	Low	Agree
Key: BM, bone marrow; ECOG, Eastern Cooperative Oncology Group; IDMC, independent data monitoring committee; IPI, International Prognostic Index; LDH, lactate dehydrogenase.			

Table 5.5: Quality assessment of the European MCL Elderly trial
(Based on Appendix 7, Table 23 of the CS[2])

Study question	How is the question addressed in the study?	Risk of bias	ERG comment
Was randomisation carried out appropriately?	Randomisation performed by permutated blocks method and stratified according to study group, age and IPI, thus presumed adequate.	Low	Unclear method of generating randomisation.
Was the concealment of treatment allocation adequate?	Randomisations performed at a central data centre or the local data centre of the national study group.	Low	Agree
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Patient demographics well balanced; prognostic factors reported including age, disease stage, ECOG, LDH and BM involvement similarly distributed.	Low	Agree
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	No. Open-label study and outcomes assessed by on site investigators only. Risk of assessment bias high. Well documented that investigator assessed	High	Agree

Study question	How is the question addressed in the study?	Risk of bias	ERG comment
the risk of bias (for each outcome)?	response is less conservative than blinded response assessment.		
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	More patients discontinued in the R-FC group due to toxicity and thus did not go onto receive maintenance therapy. Evaluable patient numbers comparable across treatment groups for induction therapy. Survival estimates for induction therapy cannot be disaggregated.	Moderate	Agree with overall risk.
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No.	Low	Agree
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Primary analysis included randomised patients who had started treatment in accordance with the randomisation results. Secondary analyses were performed according to the intention-to-treat principle with standard censoring methods used to account for missing data.	Low	Agree
Was statistical powering such to detect a significant difference between treatment groups?	Yes.	Low	Agree
Key: BM, bone marrow; ECOG, Eastern Cooperative Oncology Group; IPI, International Prognostic Index; LDH, lactate dehydrogenase			

Table 5.6: Quality assessment of the StiL NHL1 trial
(Based on Appendix 7, Table 24 of the CS[2])

Study question	How is the question addressed in the study?	Risk of bias	ERG comment

Study question	How is the question addressed in the study?	Risk of bias	ERG comment
Was randomisation carried out appropriately?	Randomisation performed according to a prespecified randomisation list and stratified according to lymphoma subtype.	Low	Agree
Was the concealment of treatment allocation adequate?	Randomisations performed centrally.	Low	Agree
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Not possible to assess for MCL subgroup of interest as only age at baseline reported for these patients.	Unclear	Agree
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. Patients, treating physicians, and individuals assessing outcomes and analysing data were not masked to treatment allocation. Risk of assessment bias high. Well documented that investigator assessed response is less conservative than blinded response assessment.	High	Agree
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Not possible to assess for MCL subgroup of interest as patient flow not reported for these patients.	Unclear	Agree
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No.	Low	Agree
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	No. Analysis does however appear to have been conducted on the per-protocol population and censoring methods appear to have been used to account for missing data but not	Unclear	Agree

Study question	How is the question addressed in the study?	Risk of bias	ERG comment
	clearly defined.		
Was statistical powering such to detect a significant difference between treatment groups?	Not for the MCL subgroup of interest.	High	Agree
Key: MCL, mantle cell lymphoma; NHL, non-Hodgkin's lymphoma; StiL, Study Group of indolent Lymphomas.			

ERG comment: The ERG disagreed with the overall risk of bias for LYM-3002 and changed the rating from low' to 'high' for blinding, since this is an open trial. There were further minor disagreements highlighted in Tables 5.4 to 5.6.

5.1.5 Evidence synthesis

Page 64 of the CS states that *"Meta-analysis was not conducted as a single RCT provides the evidence base supporting the use of VR-CAP for the treatment of adult patients with previously untreated MCL unsuitable for HSCT."*

ERG comment: As only a single randomised controlled study assessing VR-CAP, the LYM-3002 trial[9], was identified, it is justified that no meta-analysis was undertaken.

5.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

Only a single RCT assessing the technology of interest, the LYM-3002 trial[9], was identified which is discussed below.

5.2.1 Study characteristics of LYM-3002

According to the CS[2], *"the systematic literature review identified one RCT investigating the clinical efficacy and safety of VR-CAP in adult patients with previously untreated MCL"*. This RCT, LYM-3002, *"directly compares VR-CAP with R-CHOP induction therapy"*. The primary study reference is a journal publication[9] while a clinical study report (CSR)[21] and two conference abstracts[19, 20] are also available. As noted in Section 5.1.3, other relevant references were incorrectly excluded.[32-34]

The trial methodology is summarised in Table 5.7 below.

Table 5.7: Summary of the LYM-3002 trial methodology (as presented by the company) (Table 9 on page 39-41 of the CS[2])

Trial design	Randomised, open-label, multicentre, Phase III study to compare the efficacy and safety of VR-CAP to that of R-CHOP in patients with newly diagnosed MCL who are unsuitable for HSCT.
Location	128 sites including Austria (n=2), Belgium (n=8), Brazil (n=7), Canada (n=2), Chile (n=1), Columbia (n=2), Czech Republic (n=3), France (n=3), Germany (n=3), Hungary (n=6), India (n=3), Israel (n=3), Italy (n=5), Japan (n=10), People's Republic of China (n=10), Poland (n=3), Portugal (n=3), Republic of Korea (n=2), Romania (n=5), Russia (n=21), Singapore (n=1), Spain (n=6), Taiwan (n=1), Thailand (n=4), Tunisia (n=2), Turkey (n=2), Ukraine (n=7) and the United States (n=3).
Inclusion criteria	Male or female, 18 years or older; MCL Stage II, III or IV; at least 1 measurable site of disease; no prior therapies for MCL; ineligible for bone marrow transplantation; ECOG PS score of ≤2; absolute neutrophil count ≥1500 cells/µL; platelet count ≥100,000 cells/µL or ≥75,000 cells/µL if thrombocytopenia secondary to MCL was present; ALT and AST ≤3 times the ULN and total bilirubin ≤1.5 times the ULN; calculated creatinine clearance ≥20 mL/min; female subjects to be post-menopausal for at least 1 year, surgically sterile or using an effective method of birth control; males to use acceptable method of contraception; signed informed consent form.
Exclusion criteria	Prior treatment with bortezomib or prior treatment for MCL; major surgery within 2 weeks prior to randomisation; peripheral neuropathy or neuropathic pain of Grade 2 or higher; diagnosed or treated for a malignancy other than MCL within 1 year of randomisation; active systemic infection requiring treatment, known diagnosis of IV or active hepatitis B; history of allergic reaction attributable to compounds containing boron, mannitol or hydroxybenzoates; known anaphylaxis or immunoglobulin E-mediated hypersensitivity to murine proteins or to any component of rituximab; subjects who would not agree to using adequate contraception; serious medical condition or psychiatric illness likely to interfere with participation in the study; concurrent treatment with another investigational agent.
Intervention	Rituximab 375 mg/m^2 IV on day 1; cyclophosphamide 750 mg/m^2 IV on day 1; doxorubicin 50 mg/m^2 IV on day 1; bortezomib 1.3 mg/m^2 IV on days 1, 4, 8 and 11; and prednisone 100 mg/m^2 orally on day 1 through day 5 of a 21-day (3-week) cycle.
Comparator	Rituximab 375 mg/m² IV on day 1; cyclophosphamide 750 mg/m² IV on day 1; doxorubicin 50 mg/m² on day 1; vincristine 1.4 mg/m² (max. total of 2 mg) IV on day 1 and prednisone 100 mg/m² orally on day 1 through day 5 of a 21 day (3-week) cycle.
Subsequent therapy	After PD was established, subsequent therapy was at the investigator's discretion.
Permitted and disallowed concomitant medication	All concomitant medications for medical conditions other than MCL were permitted as clinically indicated, as were supportive therapies other than anti-cancer treatment. Any antineoplastic agent other than study drugs (with the exception of medications that may have antineoplastic activity but are taken for other reasons), any experimental agent other than that defined in the protocol and radiation therapy were prohibited at all times during the study. Short courses of steroids (maximum of 10 days and not exceeding 100 mg/day of prednisone or equivalent) were allowed to treat symptoms in subjects with advanced disease during screening and prior to randomisation.
Primary outcome	– PFS[a], defined as the interval between the date of randomisation and

	the date of PD or death, whichever occurred first.
Secondary outcomes	– TTP[a], defined as the duration from the date of randomisation until the date of first documented evidence of PD or relapse for subjects who experienced CR or CRu; – TTNT[a], measured from the date of initiation of study treatment as per protocol (PP) to the start date of new anti-lymphoma treatment; – Duration of TFI, measured from date of last dose plus 1 day to start date of the new treatment; – ORR[a], defined as the proportion of subjects who achieved CR, CRu, or PR relative to the response-evaluable population; – CR rate, defined as the proportion of subjects who achieved CR or CRu relative to the response-evaluable population; – TTR, measured from date of randomization to date of initial response; – Duration of radiological response[a], calculated from date of initial documentation of response to date of first documented evidence of PD or death due to PD; – Duration of CR, calculated from date of initial documentation of CR to date of first documented evidence of PD or death due to PD; – Rate of durable response, defined as the proportion of patients demonstrating a response duration ≥6 months; – OS[a], measured from the date of randomisation to the date of death. – Safety evaluations[a].
Exploratory outcomes	– PRO[a], utilising the EORTC-QLQ-C30, the EQ-5D (5 level and VAS) and the BFI, as summarised in Appendix 5:Table 20; – Medical resource utilisation; – Biomarker analyses.
Efficacy evaluations	Efficacy was assessed by blinded central review of radiology by the IRC. Radiological images were collected at baseline; cycles 2, 4, 6 (and 8, if applicable); end-of-treatment; followed by every 6 weeks for 18 weeks and thereafter every 8 weeks until confirmed PD, initiation of alternate therapy, subject withdrawal, or death. Tumour assessments were also performed locally by investigators based on radiological images and lab data (e.g. bone marrow samples and serum LDH levels). The response criteria used to assess efficacy were based on modified IWRC[CS reference 65] that are summarised in Appendix 5:Table 19.
Pre-planned subgroups	IPI risk, sex, race, region, age, stage of disease at diagnosis, LDH, WBC, ECOG, renal function, liver function.

Key: ALT, alanine transaminase; AST, aspartate transaminase; BFI, Brief Fatigue Inventory; CR, complete response; CRu, complete response unconfirmed; ECOG, Eastern Cooperative Oncology Group; EORTC-QLQ-C30, European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire; EQ-5D, EuroQol-5 Dimension Questionnaire; HSCT, haematopoietic stem cell transplantation; IPI, International Prognostic Index; IRC, independent review committee; IV, intravenous; IWRC, International Workshop Response Criteria; LDH, lactate dehydrogenase; MCL, mantle cell lymphoma; ORR, overall response rate; OS, overall survival; PD, progressive disease; PFS, progression-free survival; PR, partial response; PRO, patient reported outcome; PS, performance status; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisone; TFI, treatment free interval; TTNT, time to next treatment; TTP, time to progression; TTR, time to response; ULN, upper limit of normal; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisone; WBC, white blood cell count.
Notes: [a], listed in the final appraisal scope and included in the decision problem.
Source: LYM-3002 CSR[21], Robak et al. 2015.[9]

ERG comment:

- Trial design: The design used in this in this study is in line with expectations. As noted in Section 5.1.4, the trial did not blind participants or care providers.
- Location: In response to the clarification letter[16], the company confirmed that no patients from the UK were included. Roughly a third of the participants were recruited in the European Union and North America while the other two thirds were from the "rest of the world", especially Russia and China. Given the different prevalence depending on the geographic region and potential differences in clinical standards (e.g. concomitant care), the question of generalisability arises.[35, 36] Table 5.8 lists patients included in LYM-3002 by region and country.

Table 5.8: LYM-3002 – list of patient enrolment by region in country (as presented by the company)
(Table 8 on page 70 of the CSR[21])

Table 8: Subject Enrollment by Region and Country; Intent-to-treat Analysis Set (Study 26866138-LYM-3002)

	R-CHOP	VcR-CAP	Total
Analysis set: intent-to-treat subjects	244	243	487
Region/country			
European Union	69 (28.3%)	67 (27.6%)	136 (27.9%)
Belgium	10 (4.1%)	16 (6.6%)	26 (5.3%)
Poland	9 (3.7%)	10 (4.1%)	19 (3.9%)
Czech Republic	5 (2.0%)	9 (3.7%)	14 (2.9%)
Spain	8 (3.3%)	6 (2.5%)	14 (2.9%)
Hungary	5 (2.0%)	8 (3.3%)	13 (2.7%)
Romania	9 (3.7%)	4 (1.6%)	13 (2.7%)
Austria	6 (2.5%)	4 (1.6%)	10 (2.1%)
Italy	6 (2.5%)	3 (1.2%)	9 (1.8%)
Germany	7 (2.9%)	1 (0.4%)	8 (1.6%)
Portugal	3 (1.2%)	4 (1.6%)	7 (1.4%)
France	1 (0.4%)	2 (0.8%)	3 (0.6%)
North America Region	8 (3.3%)	6 (2.5%)	14 (2.9%)
Canada	6 (2.5%)	1 (0.4%)	7 (1.4%)
United States of America	2 (0.8%)	5 (2.1%)	7 (1.4%)
Rest of World	167 (68.4%)	170 (70.0%)	337 (69.2%)
Russia	57 (23.4%)	42 (17.3%)	99 (20.3%)
People's Republic of China	34 (13.9%)	61 (25.1%)	95 (19.5%)
Ukraine	18 (7.4%)	16 (6.6%)	34 (7.0%)
Brazil	9 (3.7%)	13 (5.3%)	22 (4.5%)
Thailand	9 (3.7%)	10 (4.1%)	19 (3.9%)
Japan	11 (4.5%)	7 (2.9%)	18 (3.7%)
India	9 (3.7%)	3 (1.2%)	12 (2.5%)
Israel	5 (2.0%)	2 (0.8%)	7 (1.4%)
Tunisia	3 (1.2%)	3 (1.2%)	6 (1.2%)
Turkey	5 (2.0%)	1 (0.4%)	6 (1.2%)
Colombia	2 (0.8%)	3 (1.2%)	5 (1.0%)
Republic of Korea	2 (0.8%)	3 (1.2%)	5 (1.0%)
Chile	1 (0.4%)	2 (0.8%)	3 (0.6%)
Singapore	1 (0.4%)	2 (0.8%)	3 (0.6%)
Taiwan	1 (0.4%)	2 (0.8%)	3 (0.6%)

Key: R-CHOP=rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone;
VcR-CAP=VELCADE, rituximab, cyclophosphamide, doxorubicin, and prednisone.
Note: Percentages calculated with the number of subjects in each group as denominator.

[TSUB01A.rtf] [JNJ-26866138\LYM3002\DBR_CORE\RE_CSR\tsub01a.sas] 24FEB2014, 16:11

- Inclusion and exclusion criteria:
  - The inclusion criteria are narrower than those defined in the final scope[1] as the criteria used in the CS specified performance scores and laboratory thresholds. Neither the CS nor the CSR give a justification; however, these criteria are

likely to be clinically justified. However, it should be noted that patients who might be suitable according to the final scope *("People with previously untreated mantle cell lymphoma, who are not going to have a stem cell transplant"*)[1] might not have been eligible for inclusion in the LYM-3002 trial.

- The final scope[1] uses two different definitions of the population of interest in regards to the haematopoietic stem cell transplantations (HSCT): 1) "People with previously untreated mantle cell lymphoma, *who are not going to have a stem cell transplant*" and 2) "adult patients with previously untreated mantle cell lymphoma *who are unsuitable for haematopoietic stem cell transplantation*". With that in mind, it should be noted that following a protocol amendment the LYM-3002 trial enrolled "patients ineligible *or not considered* for HSCT".[2] This amendment can be considered to be in line with the first but not second definition used in the final scope.

  Later, *"concerns over the heterogeneity and interpretability of the study results resulted in a further amendment, realigning to the original criteria, and only patients who were not eligible for HSCT as assessed by the treating physician, that is, patients considered medically ineligible (e.g. due to age or the presence of co-morbid conditions that may have a negative impact on the tolerability to transplantation), were subsequently enrolled"*.[2] This second amendment could be considered to be in line with both the first and second definition used in the final scope.

  The trial included 80 patients (16.4%; 38/243 in VR-CAP arm, 42/244 in R-CHOP arm) *"were suitable for HSCT from a medical perspective but availability and/or socio-economic reasons prevented access"*.

  The ERG sought clarification of how these patients were accounted for. The company responded that these patients were balanced across the two patient groups and that post-hoc analysis revealed consistent treatment effect regardless of eligibility for HSCT.[16] The ERG found these statements to be true for overall survival but not progression-free survival. Progression-free survival was significantly favourable for VR-CAP for those patients that were medically ineligible for HSCT but non-significantly favourable for VR-CAP for patients who were medically eligible. The difference in statistical significance is likely due to the reduced patient numbers who were medically eligible (n=80 versus n=407). Overall the ERG does not think the difference in population inclusion criteria will significantly influence the overall results.

- Intervention: The ERG does not have any specific comments on this aspect.
- Comparator: The ERG does not have any specific comments on this aspect.
- Subsequent therapy and permitted and disallowed concomitant medication: Subsequent or concomitant were permitted at the investigator's discretion or as clinically indicated. This is in line with clinical practice.

- Primary, secondary and exploratory outcomes: The outcomes described in the final scope were as follows: *"overall survival, progression-free survival, overall response rate, duration of response/remission, time to new anti-lymphoma treatment/time to progression, adverse effects of treatment, health-related quality of life"*.[1] In the LYM-3002 trial, progression-free survival is the primary outcome. The secondary outcomes used in the LYM-3002 trial cover most of the other outcomes defined in the final scope. Additionally, some other measures of response are reported, e.g. time to response (TTR), duration of radiological response or duration of complete response. Patient reported outcomes, medical resource utilisation and biomarker analyses were reported as exploratory analyses.
- Efficacy evaluations: The ERG has no specific comments on this aspect.
- Pre-planned subgroups: As reported in the table above, 11 subgroups were pre-planned. For "region", three regional subgroups (European Union, North America region and rest of the world) were pre-specified. Due to small numbers of participants in the North American region, Europe and North America were combined post hoc while the North American and Western European subgroup was pre-specified. As discussed before, only approximately a third of the participants of LYM-3002 were recruited in Northern America or Western Europe. Results for this subgroup will be presented and discussed alongside the results for the whole group.

Length of follow-up

"Subjects received 6 to 8 cycles (18 to 24 weeks) of treatment depending upon the response documented at the cycle 6 assessment. The total study duration from randomisation of the first patient until the last PFS event required for the final analysis was expected to be approximately 42 months (24 months for enrolment and 18 months for follow-up)".[2]

"There are no additional interim analyses available between that provided in the submission (Dec 2013 cut-off) and the final analysis planned in 2017. There were 3 interim analysis conducted prior to the primary analysis (Dec 2013 cut-off) in 2009, 2010 and 2011; these do not provide additional clinical data".[16]

ERG comment: The length of follow-up seems adequate. However, it should be noted that the LYM-3002 trial is ongoing and only results of an interim analysis were reported in the CS.

Statistical analysis

Statistical analyses in the LYM-3002 are detailed in Table 5.9.

Table 5.9: LYM-3002 – summary of statistical analyses
(Table 10 on page 42 of the CS[2])

Hypothesis objective	VR-CAP provides greater benefit in newly diagnosed MCL patients with Stage II, III or IV disease compared to R-CHOP, as assessed by significant prolongation of PFS.
Statistical analysis	All statistical tests were 2-sided. The primary hypothesis was tested at the 0.05 significance level (overall). The significance level at the interim analysis was determined by the observed number of events at the time of the interim analysis per

	the O'Brien-Fleming spending function. The secondary hypotheses were tested sequentially at the nominal 0.05 significance level in the following order: 1) TTP; 2) TTNT; 3) CR rate; and 4) OS. A secondary hypothesis was tested only if the primary hypothesis was rejected along with all the secondary hypotheses that preceded it. However, OS was to be analysed regardless of what happened to the other endpoints in the hierarchical test. The KM method was used to estimate the distribution of overall PFS for each treatment group. The primary treatment comparison was based on a stratified log-rank test. The HR and its 95% CI were estimated based on a stratified Cox's model, with treatment as the explanatory variable. Stratification factors were IPI (0-1, 2, 3, and 4-5) and stage of disease at diagnosis (II, III, and IV). The same methodology was used for other time to event endpoints, except duration of response and duration of complete response where the KM method was used as a descriptive summary. For response-related endpoints, the stratified Cochran–Mantel–Haenszel test was used to test treatment difference. The Mantel–Haenszel estimate of the OR and its 95% CI were also calculated.
Sample size, power calculation	Assuming that treatment with VR-CAP would improve the median PFS by 40% (from 18 to 25 months), a total number of 295 events would provide 80% power (alpha=0.05, 2-sided) to detect such improvement. Assuming a 24-month accrual and an 18-month follow-up, a total of 486 subjects were needed for the study (243 subjects per treatment group).
Data management, patient withdrawals	The primary efficacy analysis set was the ITT population, which included all randomised subjects. The secondary efficacy analysis set included (1) the PP population, which included all randomised subjects who met eligibility criteria, received at least 1 dose of study drug and underwent at least 1 post-baseline disease assessment, and (2) the response-evaluable population, which included all randomised subjects who received at least 1 dose of study drug, had at least 1 measurable tumour mass at baseline and had at least 1 post-baseline assessment. The safety population was defined as all randomised subjects who received at least 1 dose of the study drug. Subjects may have been withdrawn from the study for any of the following reasons: lost to follow-up, withdrawal of consent, death. If a subject withdrew before the end of the treatment phase, end-of-treatment and follow-up assessments were obtained. In efficacy analyses, subjects who withdrew from the study were censored at the time of the last adequate (CT scan available) disease assessment.
Key: CI, confidence interval; CT, computed tomography; CR, complete response; HR, hazard ratio; IPI, International Prognostic Index; ITT, intent-to-treat; KM, Kaplan-Meier; MCL, mantle cell lymphoma; OR, odds ratio; OS, overall survival; PFS, progression-free survival; PP, per protocol; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisone; TTNT, time to next treatment; TTP, time to progression; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisone. **Source:** LYM-3002 CSR[21], Robak et al. 2015.[9]	

ERG comment: The ERG does not have any specific comments on this section.

Subgroup analyses

ERG comment: Eleven subgroups were pre-planned. As discussed before, only approximately a third of the participants of LYM-3002 were recruited in Northern America or Western Europe. Results for this subgroup will be presented and discussed alongside the results for the whole group.

Sample size, power calculation

LYM-3002 was powered for the primary endpoint 'progression-free survival' (PFS).[2] In reply to clarification letter, the company stated that *"no power calculations were made for any of the other endpoints"*.

ERG comment: It should be noted that endpoints other than PFS might have been underpowered. When considering only participants from Europe and North America (cf. Section 5.2.1), all outcomes are likely to be underpowered.

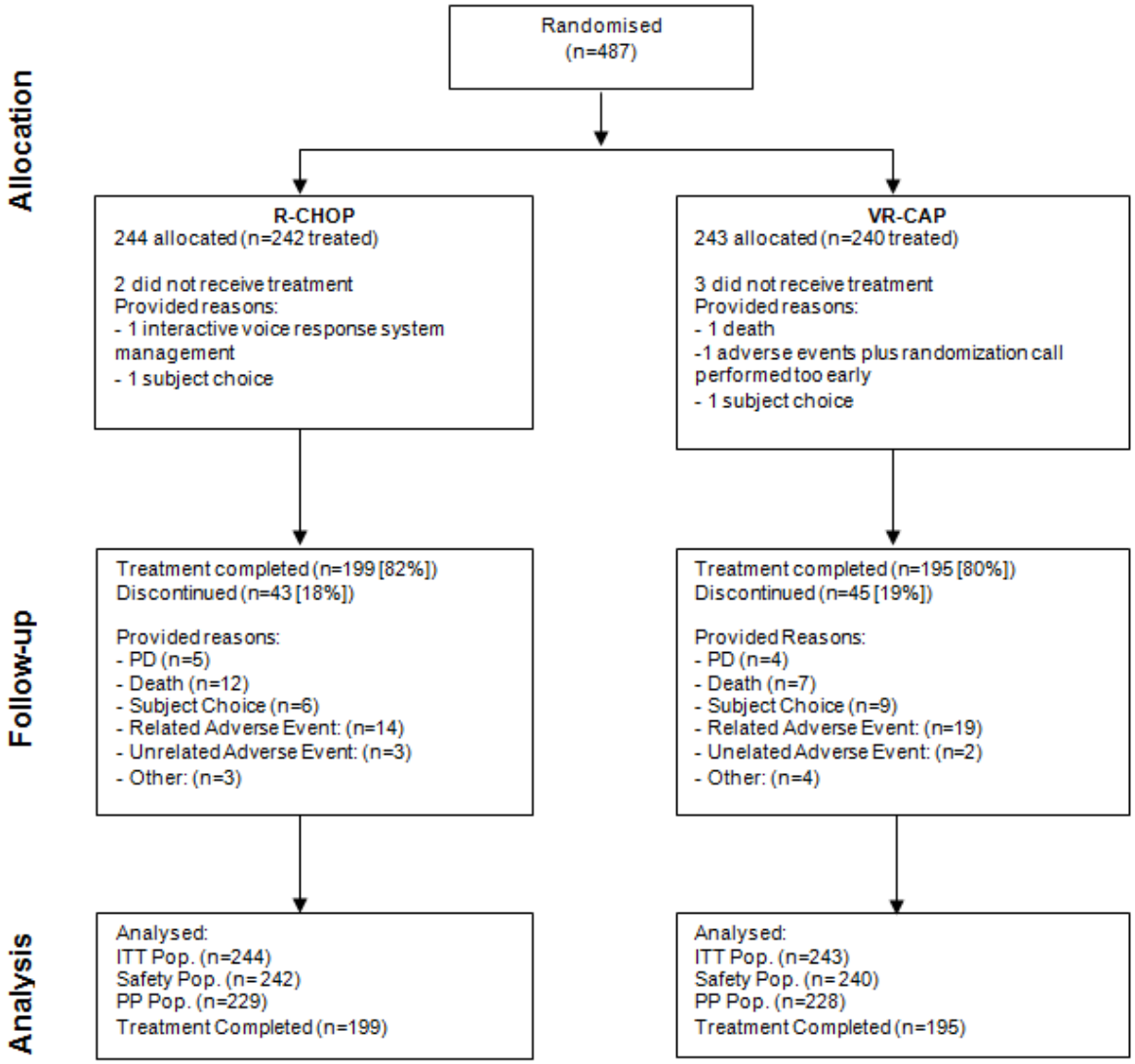
Discontinuation and censoring

ERG comment: The ERG agrees with this approach to account for missing data.

Eligibility criteria

The flow of patients through the study is presented in Figure 5.1.

Figure 5.1: LYM-3002 – study flow diagram
(Table 10 on page 42 of the CS[2])



Key: ITT, intent-to-treat; PD, progressive disease; PP, per protocol; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisone; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisone.
Source: LYM-3002 CSR[21], Robak et al. 2015.[9]

ERG comment: Numbers of discontinuations and participants analysed were comparable between the groups.

5.2.2 Results

As detailed in Section 5.1.5, only a single RCT assessing the technology of interest, the LYM-3002 trial[9], was identified.

According to Section 4.7 of the company submission[2], *"the clinical cut-off date for the primary analysis presented in this submission was 2 December 2013"* and *"the final analysis is expected in 2017"*. In response to the clarification letter,[16] the company stated that *"there are no additional interim analyses available between that provided in the submission (Dec*

2013 cut-off) and the final analysis planned in 2017. There were 3 interim analysis conducted prior to the primary analysis (Dec 2013 cut-off) in 2009, 2010 and 2011".

The final scope[37] lists the following outcome measures:

- Overall survival
- Progression-free survival
- Overall response rate
- Duration of response/remission
- Time to new anti-lymphoma treatment/time to progression
- Adverse effects of treatment
- Health-related quality of life

Results for these outcomes presented in the CS are discussed below. Where available, results for all participants are presented alongside the subgroup analysis for participants from Western Europe or North America (cf. Section 5.2.1).

Overall survival

According to the CS[2], *"overall survival data are not yet mature in the LYM-3002 trial. However, interim analysis indicates a consistent survival trend favouring the VR-CAP group"* (see table 5.10).

Table 5.10: LYM-3002 – overall survival (ITT set)
(Table 19 on page 57 of the CS[2])

	VR-CAP (n=243)	R-CHOP (n=244)
OS, median days (95% CI)	NE (1704.0; NE)	1714.0 (1436.0; NE)
HR (95% CI)	0.80 (0.59; 1.10)	
p-value	0.173	
18-month survival rate % (95% CI)	84.9 (79.6; 88.9)	83.8 (78.4; 87.9)
3-year survival rate % (95% CI)	72.2 (65.6; 77.8)	67.9 (61.1; 73.8)
4-year survival rate % (95% CI)	64.4 (56.4; 71.4)	53.9 (45.2; 61.9)

Key: CI, confidence interval; HR, hazard ratio; ITT, intent-to-treat; NE, not estimable; OS, overall survival; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisone; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisone.
Source: LYM-3002 CSR[21], Robak et al. 2015.[9]

In the subgroup of Western European and North American participants, *"twenty-nine events were observed, 15 R-CHOP subjects (31%) and 14 VcR-CAP subjects (33%). Median OS was not reached in either group (...) The median OS was not reached in either group, although there was a trend favouring VcR-CAP (HR=0.83; 95% CI: 0.34, 2.02) consistent with the overall ITT population"*[21] (Table 5.11).

Table 5.11: LYM-3002 – overall survival (ITT set) – Western European and North American participants
(Table 50 on page 137 of the CSR[21])

	R-CHOP	VcR-CAP
Analysis set: intent-to-treat subjects from US/Canada/Western European countries	49	42
Descriptive[a]		
Overall survival (days)		
Number of assessed	49	42
Number of censored (%)	34 (69.4%)	28 (66.7%)
Number of events (%)	15 (30.6%)	14 (33.3%)
25% quantile (95% CI)	896.0 (464.0;1466.0)	682.0 (173.0;NE)
Median (95% CI)	NE (1436.0;NE)	NE (786.0;NE)
75% quantile (95% CI)	NE (NE ;NE)	NE (NE ;NE)
P-value[b]	0.839	
Hazard ratio (95% CI)[c]		1.09 (0.47; 2.52)
18-month survival rate % (95% CI)	85.2 (71.4; 92.6)	78.2 (62.2; 88.0)
3-year survival rate % (95% CI)	69.8 (53.0; 81.6)	63.8 (46.4; 76.9)
4-year survival rate % (95% CI)	62.1 (40.2; 77.9)	63.8 (46.4; 76.9)

Key: R-CHOP=rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone;
VcR-CAP=VELCADE, rituximab, cyclophosphamide, doxorubicin, and prednisone.
[a] Based on Kaplan-Meier product limit estimates.
[b] Based on Log rank test stratified with IPI risk and stage of disease.
[c] Hazards ratio estimate is based on a Cox´s model stratified by IPI risk and stage of disease. A hazard ratio < 1 indicates an advantage for VcR-CAP.
NE: Not estimable.
Western European countries include: Austria, Belgium, France, Germany, Italy, Portugal and Spain.

ERG comment: In the overall group, no statistically significant difference between VR-CAP and R-CHOP regarding overall survival was seen (HR 0.80, 95% CI 0.59 to 1.10). In the subgroup of Western European and North American participants, the results also showed no statistically significant difference (HR 1.09, 95% CI 0.47 to 2.52). As indicated by the wide confidence, numbers of included participants as well as events were small. In addition, the study might not have been adequately powered for this endpoint (see Section 5.2.1). Therefore, results should be interpreted with caution. It should be noted that the "*overall survival data are not yet mature in the LYM-3002 trial*",[2] more comprehensive results will be available with the final analysis in 2017.

Progression-free survival

"The clinical cut-off date for the primary analysis presented in this submission was 2 December 2013.[CS references 59+60] *The final analysis is expected in 2017 (...)In the primary analysis of PFS based on IRC assessment of the ITT population, median PFS was 751 days (24.7 months) in subjects randomised to VR-CAP compared with 437 days (14.4 months) in subjects randomised to R-CHOP (HR=0.63, p<0.001)"*[2] (Table 5.12).

"The results for the IRC assessment of PFS for the total ITT population and the US/Canada/Western European subgroup appear consistent (HR= 0.63 and 0.77, respectively) in favor of the VcR-CAP group. Results from the subgroup of 14 subjects from US/Canada, which appeared to be an outlier, were inconsistent with other subgroup analyses. The separate analysis of the Western European subgroup (HR=0.49; 95% CI: 0.25, 0.95) appears more consistent with the overall study results"[21] (Table 5.12).

Table 5.12: LYM-3002 – progression-free survival based on IRC assessment, ITT analysis set
(Table 47 on page 134 of the CSR[21])

	R-CHOP (N=244)	VcR-CAP (N=243)		
	median PFS (months)		HR[a] (95% CI)	p-value[b]
All Subjects (ITT)	N=244	N=243		
IRC assessment	437.0 (14.4)	751.0 (24.7)	0.63 (0.50, 0.79)	0.001
Investigator assessment	490.0 (16.1)	934.0 (30.7)	0.51 (0.41, 0.65)	0.001
US/Canada/Western Europe Combined	N=49	N=42		
IRC assessment	437.0 (14.4)	592.0 (19.4)	0.77 (0.43, 1.38)	0.380
Investigator assessment	561.0 (16.1)	855.0 (28.1)	0.59 (0.32, 1.09)	0.090
US/Canada Alone	N=8	N=6		
IRC assessment	782.0 (25.7)	288.5 (9.5)	4.31 (0.44; 41.96)	0.171
Investigator assessment	895.0 (29.4)	529.0 (17.4)	3.77 (0.32; 43.72)	0.259
Western Europe Alone	N=41	N=36		
IRC assessment	393.0 (12.9)	604.0 (19.8)	0.49 (0.25, 0.95)	0.031
Investigator assessment	527.0 (17.3)	1,009 (33.1)	0.43 (0.22, 0.82)	0.010

Key: R-CHOP=rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone;
VcR-CAP=VELCADE, rituximab, cyclophosphamide, doxorubicin, and prednisone.
PFS = progression-free survival; IRC= Independent Radiology Review Committee; INV=investigator
[a] Hazards ratio estimate is based on a Cox´s model stratified by IPI risk and stage of disease. A hazard ratio <1 indicates an advantage for VcR-CAP.
[b] Based on Log rank test stratified with IPI risk and stage of disease.
Western European countries include: Austria, Belgium, France, Germany, Italy, Portugal, and Spain

ERG comment: In the overall group, a statistically significant difference between VR-CAP and R-CHOP regarding progression-free survival was seen (HR 0.63, 95% CI 0.50 to 0.79, in favour of VR-CAP). In the subgroup of Western European and North American participants, the results did not show a statistically significant difference (HR 0.77, 95% CI 0.43 to 1.38). For the Western European Group, the HR was 0.49 (95-CI 0.25 to 0.95, in favour of VR-CAP). As indicated by the wide confidence, numbers of included participants as well as events were small. Therefore, results should be interpreted with caution. It should be noted that more comprehensive results will be available with the final analysis in 2017.

Overall response rate

"Both VR-CAP and R-CHOP were associated with a high rate of clinical response. However, VR-CAP demonstrated a superior depth of response with a higher rate of overall CR, shorter TTR and longer DoR. Based on IRC assessment of the response evaluable population, 53.3% of subjects in the VR-CAP group demonstrated a complete response (CR + CRu) to treatment (confirmed by evidence of bone marrow clearance and LDH normalisation), compared with 41.7% of subjects in the R-CHOP group (OR: 1.69; p=0.007). Investigator response assessment demonstrated consistent results in favour of the VR-CAP group (41.5% vs 27.6%; OR: 1.89; p=0.002)"[2] (Table 5.13).

Table 5.13: LYM-3002 – best overall response, response evaluable analysis set
(Table 17 on page 55 of the CS[2])

	IRC assessment		Investigator assessment	
	VR-CAP (n=229)	**R-CHOP** (n=228)	**VR-CAP** (n=229)	**R-CHOP** (n=228)
Overall response (CR+CRu+PR), n (%)	211 (92.1)	204 (89.5)	219 (95.6)	209 (91.7)
OR (95% CI)	1.428 (0.749, 2.722)		2.022 (0.920, 4.446)	
p-value	0.275		0.073	
Overall complete response, n (%)	122 (53.3)	95 (41.7)	95 (41.5)	63 (27.6)
OR (95% CI)	1.688 (1.148, 2.481)		1.884 (1.259, 2.819)	
p-value	0.007		0.002	

Key: CI, confidence interval; CR, complete response; CRu, complete response unconfirmed; IRC, independent review committee; LDH, lactate dehydrogenase; OR, odds ratio; PR, partial response; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisone; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisone.
Notes: [a], confirmed by evidence of bone marrow clearance and LDH normalisation.
Source: LYM-3002 CSR[21], Robak et al. 2015.[9]

"Overall complete response, overall radiological complete response, and overall response rate by IRC were slightly higher in the R-CHOP group compared with the VcR-CAP group (...) Overall complete response was 47.8% in the R-CHOP group versus 45.0% in the VcR-CAP group (OR=1.150; p=0.763).[sic!] *Overall radiological complete response was 76.1% in the R-CHOP group versus 72.5% in the VcR-CAP group (OR=1.094; p=0.872). Overall response rate was 89.1% in the R-CHOP group and 85.0% in the VcR-CAP group (OR=1.163; p=0.811)"*[21] (Table 5.14).

Table 5.14: LYM-3002 – best overall response – Western European and North American participants
(Table 49 on page 136 of the CSR[21])

	R-CHOP	VcR-CAP	P-value[c]	OR[d](95% CI)
Analysis set: response-evaluable subjects from US/Canada/Western European countries	46	40		
Overall complete response (CR+CRu)	25 (54.3%)	20 (50.0%)	0.965	1.020(0.420, 2.476)
CR	22 (47.8%)	18 (45.0%)		
CRu	3 (6.5%)	2 (5.0%)		
Overall radiological complete response (CR+CRu)[a]	35 (76.1%)	29 (72.5%)	0.872	1.094(0.374, 3.200)
Radiological CR	32 (69.6%)	27 (67.5%)		
Radiological CRu	3 (6.5%)	2 (5.0%)		
Partial response (PR)	6 (13.0%)	5 (12.5%)		
Overall response (CR+CRu+PR)[b]	41 (89.1%)	34 (85.0%)	0.811	1.163(0.318, 4.248)
Stable disease (SD)	1 (2.2%)	0		
Progressive disease (PD)	4 (8.7%)	6 (15.0%)		

Key: R-CHOP=rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone; VcR-CAP=VELCADE, rituximab, cyclophosphamide, doxorubicin, and prednisone; OR=Odds Ratio; CR=Complete Response; PR=Partial Response; SD=Stable Disease; PD=Progressive Disease.
[a] Include all radiological CR+CRu, regardless the verification by bone marrow and LDH.
[b] Include all radiological CR+CRu+PR, regardless the verification by bone marrow and LDH.
[c] P-value from the Cochran Mantel-Haenszel Chi-Squared test, with IPI and Stage of Disease as stratification factors.
[d] Mantel-Haenszel estimate of the common odds ratio for stratified tables is used, with IPI and Stage of Disease as stratification factors. An odds ratio (OR) > 1 indicates an advantage for VcR-CAP.
Percentage calculated with the number of subjects in each group as denominator.
Western European countries include: Austria, Belgium, France, Germany, Italy, Portugal and Spain.

ERG comment: In the overall group, no statistically significant difference between VR-CAP and R-CHOP regarding overall response was seen (OR 1.43, 95% CI 0.75 to 2.72). In the subgroup of Western European and North American participants, the results did not show a statistically significant difference (OR 1.16, 95% CI 0.32 to 4.25). As indicated by the wide confidence, numbers of included participants as well as events were small. Therefore, results should be interpreted with caution. It should be noted that more comprehensive results will be available with the final analysis in 2017.

Duration of response/remission

"Median DoR was markedly longer in subjects responding to VR-CAP treatment compared with subjects responding to R-CHOP treatment (...) Based on IRC assessment, median duration of overall response, median DoR for complete responders and median duration of complete response were more than twice as long in the VR-CAP treatment group compared with the R-CHOP treatment group: 1,110 days (36.5 months) versus 459 days (15.1 months); 1,282 days (42.1 months) versus 563 days (18.5 months); and 1,282 days (42.1 months) versus 547 days (18.0 months), respectively"[2] (Table 5.15).

Table 5.15: LYM-3002 – time to response and duration of response, response evaluable analysis set

(Table 18 on page 56 of the CS[2])

	IRC assessment		**Investigator assessment**	
	VR-CAP (n=229)	**R-CHOP** (n=228)	**VR-CAP** (n=229)	**R-CHOP** (n=228)
Median TTR, days (95% CI)	42.0 (42.0; 43.0)	50.0 (44.0; 74.0)	43.0 (42,0; 43.0)	48.0 (43.0; 75.0)
HR (95% CI)	1.54 (1.26; 1.89)		1.44 (1.18; 1.75)	
p-value	<0.001		<0.001	
Median DoR, days (95% CI)	1110.0 (813.0; 1420.0)	459.0 (379.0; 518.0)	1060.0 (892.0; 1366.0)	491.0 (410.0; 560.0)
DoR for complete responder, days (95% CI)	1282.0 (933.0; 1602.0)	563.0 (486.0; 738.0)	NE (1255.0; NE)	684.0 (603.0; 913.0)
Duration of CR, days (95% CI)	1282.0 (933.0; 1495.0)	547.0 (425.0; 711.0)	1516.0 (1150.0; NE)	568.0 (476.0; 833.0)
Durable overall response (CR+CRu+PR) rate, n (%)	172 (75.1)	151 (66.2)	187 (81.7)	164 (71.9)
OR (95% CI)	1.563 (1.032, 2.367)		1.693 (1.090, 2.632)	
p-value	0.035		0.018	
Durable CR (CR+CRu) rate, n (%)	112 (48.9)	82 (36.0)	90 (39.3)	57 (25.0)
Odds ratio (95% CI)	1.799 (1.219, 2.656)		1.960 (1.302, 2.950)	
p-value	0.003		0.001	

Key: CI, confidence interval; CR, complete response; CRu, complete response unconfirmed; DoR, duration of response; IRC, independent review committee; NE, not estimable; OR, odds ratio; PR, partial response; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisone; TTR, time to response; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisone.
Source: LYM-3002 CSR[21], Robak et al. 2015.[9]

ERG comment: As shown in the table above, the median duration of response in the VR-CAP group (1110 days, 95% CI 813 to 1420) is longer than in R-CHOP group (459 days, 95% CI 379 to 518). No results for the subgroup of Western European and North American

participants were available. It should be noted that more comprehensive results will be available with the final analysis in 2017.

Time to new anti-lymphoma treatment/time to progression

"*Based on IRC assessment of the ITT population, median TTP was 929 days (30.5 months) in subjects randomised to VR-CAP compared with 490 days (16.1 months) in subjects randomised to R-CHOP (HR=0.58; p<0.001)*"[2] (Table 5.16).

Table 5.16: LYM-3002 – Time to disease progression
(Table 26 on page 117 of the CS appendices[15])

	IRC assessment		Investigator assessment	
	VR-CAP (n=243)	**R-CHOP** (n=244)	**VR-CAP** (n=243)	**R-CHOP** (n=244)
Time to disease progression, days[a]				
Number of censored (%)	129 (53.1)	96 (39.3)	135 (55.6)	86 (35.2)
Number of events (%)	114 (46.9)	148 (60.7)	108 (44.4)	
25% quantile (95% CI)	363.0 (282.0; 446.0)	298.0 (248.0; 319.0)	494 (386.0; 602.0)	304.0 (271.0; 356.0)
Median (95% CI)	929.0 (696.0; 1245.0)	490.0 (417.0; 550.0)	1066.0 (862.0; 1299.0)	512.0 (449.0; 601.0)
75% quantile (95% CI)	1698.0 (1515.0; NE)	1049.0 (745.0; NE)	NE (1639.0; NE)	950.0 (734.0; NE)
p-value[b]	<0.001		<0.001	
Hazard ratio (95% CI)	0.58 (0.45; 0.74)		0.47 (0.36; 0.60)	

Key: CI, confidence interval; IPI, Internal Prognostic Index; IRC, Independent Review Committee; ITT, intent-to-treat; NE, not estimable; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisone; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisone.
Notes: [a] based on Kaplan–Meier product limit estimates; [b] based on log rank test stratified with IPI risk and stage of disease.
Source: LYM-3002 CSR.[3]

In the subgroup of participants from Western Europe and North America, "*based on the IRC, median TTP* [time to progression] *was 475 days (15.6 months) in the R-CHOP group, compared with 682 days (22.4 months) in the VcR-CAP group (HR=0.63; p=0.132) (...) TTP in the R-CHOP group was similar to that in the larger ITT population (16.1 months), while TTP in the VcR-CAP group was somewhat shorter than that in the larger population (30.5 months)*"[21] (Table 5.17).

Table 5.17: LYM-3002 – summary to disease progression, ITT set – Western European and North American participants
(Table 48 on page 135 of the CSR[21])

	R-CHOP	VcR-CAP
Analysis set: intent-to-treat subjects from US/Canada/Western European countries	49	42
Descriptive[a]		
Time to disease progression(days)		
Number of assessed	49	42
Number of censored (%)	16 (32.7%)	19 (45.2%)
Number of events (%)	33 (67.3%)	23 (54.8%)
25% quantile (95% CI)	269.0 (177.0; 379.0)	163.0 (44.0; 386.0)
Median (95% CI)	475.0 (362.0; 634.0)	682.0 (274.0;1318.0)
75% quantile (95% CI)	782.0 (534.0;NE)	1318.0 (1145.0;1718.0)
P-value[b]	0.132	
Hazard ratio (95% CI)[c]		0.63 (0.34; 1.16)

Key: R-CHOP=rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone;
VcR-CAP=VELCADE, rituximab, cyclophosphamide, doxorubicin, and prednisone.
[a] Based on Kaplan-Meier product limit estimates.
[b] Based on Log rank test stratified with IPI risk and stage of disease.
[c] Hazards ratio estimate is based on a Cox´s model stratified by IPI risk and stage of disease. A hazard ratio < 1 indicates an advantage for VcR-CAP.
NE: Not estimable.
Western European countries include: Austria, Belgium, France, Germany, Italy, Portugal and Spain.

ERG comment: In the overall group, a statistically significant difference between VR-CAP and R-CHOP regarding time to next anti-lymphoma treatment was seen (HR 0.50, 95% CI 0.30 to 0.65, in favour of VR-CAP). In the subgroup of Western European and North American participants, the results regarding disease progression did not show a statistically significant difference (HR 0.63, 95% CI 0.34 to 1.16), results for time to next anti-lymphoma treatment were not reported in this subgroup. As indicated by the wide confidence, numbers of included participants as well as events were small. Therefore, results should be interpreted with caution. It should be noted that more comprehensive results will be available with the final analysis in 2017.

Adverse events of treatment

"The safety population, defined as all randomised subjects who received at least 1 dose of study medication, consisted of 240 subjects in the VR-CAP group and 242 subjects in the R-CHOP group.

Both chemotherapy induction regimens were generally well tolerated, with low rates of discontinuation due to an AE and low fatality on treatment in both groups. Almost all subjects in both treatment groups experienced a treatment-emergent AE, although VR-CAP was associated with a slightly higher rate of Grade 3 or higher AEs and serious AEs (SAEs) (...).

The majority of the fatal study drug-related AEs (3 subjects in each treatment group) were of infectious origin, none of which were solely attributed to the bortezomib or vincristine elements of the chemotherapeutic regimens"[2] (Table 5.18).

Adverse events were also summarised by study centre geographic region (European Union, US and Canada, and rest of the world (ROW)) (Table 5.19).

Table 5.18: LYM-3002 – safety and tolerability, safety analysis set
(Table 27 on page 75 of the CS[2])

	VR-CAP (n=240)	R-CHOP (n=242)
Any treatment-emergent AE, n (%)	238 (99.2)	238 (98.3)
At least one related[a]	231 (96.3)	226 (93.4)
None related	7 (2.9)	12 (5.0)
Any serious AE, n (%)	90 (37.5)	72 (29.8)
At least one related[a]	78 (32.5)	50 (20.7)
None related	12 (5.0)	22 (9.1)
Maximum severity of any AE, n (%)	238 (99.2)	238 (98.3)
Grade 1	2 (0.8)	6 (2.5)
Grade 2	13 (5.4)	26 (10.7)
Grade 3	31 (12.9)	53 (21.9)
Grade 4	176 (73.3)	136 (56.2)
Grade 5	16 (6.7)	17 (7.0)
Treatment discontinuation due to AEs, n (%)	21 (8.8)	17 (7.0)
Related to study drug[a]	19 (7.9)	14 (5.8)
Fatality due to AEs, n (%)	16 (7.0)	17 (7.0)
Related to study drug[a]	5 (2.0)	7 (3.0)

Key: AE, adverse event; CSR, clinical study report; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisone; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisone.
Notes: [a], related to any study drug
Source: LYM-3002 CSR[21], Robak et al. 2015.[9]

Table 5.19: LYM-3002 – overview of treatment-emergent AE by region, safety analysis set
(Table 83 on page 203 of the CSR[21])

	R-CHOP				VcR-CAP			
		Region				Region		
	Total	EU	USA/CAN	ROW	Total	EU	USA/CAN	ROW
Analysis set: safety subjects	242	68	8	166	240	65	6	169
Any treatment-emergent adverse event	238 (98.3%)	66 (97.1%)	8 (100.0%)	164 (98.8%)	238 (99.2%)	64 (98.5%)	6 (100.0%)	168 (99.4%)
At least one related [a]	226 (93.4%)	66 (97.1%)	8 (100.0%)	152 (91.6%)	231 (96.3%)	62 (95.4%)	6 (100.0%)	163 (96.4%)
None related	12 (5.0%)	0	0	12 (7.2%)	7 (2.9%)	2 (3.1%)	0	5 (3.0%)
Any serious adverse event	72 (29.8%)	17 (25.0%)	0	55 (33.1%)	90 (37.5%)	33 (50.8%)	2 (33.3%)	55 (32.5%)
At least one related [a]	50 (20.7%)	11 (16.2%)	0	39 (23.5%)	78 (32.5%)	29 (44.6%)	1 (16.7%)	48 (28.4%)
None related	22 (9.1%)	6 (8.8%)	0	16 (9.6%)	12 (5.0%)	4 (6.2%)	1 (16.7%)	7 (4.1%)
Maximum severity of any AE	238 (98.3%)	66 (97.1%)	8 (100.0%)	164 (98.8%)	238 (99.2%)	64 (98.5%)	6 (100.0%)	168 (99.4%)
Grade 1	6 (2.5%)	0	1 (12.5%)	5 (3.0%)	2 (0.8%)	0	0	2 (1.2%)
Grade 2	26 (10.7%)	10 (14.7%)	1 (12.5%)	15 (9.0%)	13 (5.4%)	4 (6.2%)	0	9 (5.3%)
Grade 3	53 (21.9%)	16 (23.5%)	3 (37.5%)	34 (20.5%)	31 (12.9%)	5 (7.7%)	2 (33.3%)	24 (14.2%)
Grade 4	136 (56.2%)	38 (55.9%)	3 (37.5%)	95 (57.2%)	176 (73.3%)	49 (75.4%)	4 (66.7%)	123 (72.8%)
Grade 5	17 (7.0%)	2 (2.9%)	0	15 (9.0%)	16 (6.7%)	6 (9.2%)	0	10 (5.9%)
Treatment discontinuation due to AEs	17 (7.0%)	4 (5.9%)	0	13 (7.8%)	21 (8.8%)	7 (10.8%)	1 (16.7%)	13 (7.7%)
At least one related [a]	14 (5.8%)	4 (5.9%)	0	10 (6.0%)	19 (7.9%)	6 (9.2%)	1 (16.7%)	12 (7.1%)
None related	3 (1.2%)	0	0	3 (1.8%)	2 (0.8%)	1 (1.5%)	0	1 (0.6%)

Key: R-CHOP=rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone;
VcR-CAP=VELCADE, rituximab, cyclophosphamide, doxorubicin, and prednisone; AE=adverse event; EU=European Union;
USA/CAN=United States/Canada; ROW=Rest of World.
[a] Related to any study drug.
Note: Percentages in 'Total' column for each group calculated with the number of subjects in each group as denominator. Percentages of region sub-groups calculated with number of subjects per sub-group as denominator.
Incidence is based on the number of subjects experiencing at least one adverse event, not the number of events.
Adverse events with missing toxicity grade are not included in the table.

"In both treatment groups, the most commonly reported AEs and Grade 3 or higher AEs were haematological (blood and lymphatic system) disorders.[CS references 59+60] *There was a difference (i.e. ≥10% between treatment groups) in the incidence rates for the following AEs: neutropenia, thrombocytopenia, leukopenia, anaemia, lymphopenia, pyrexia, diarrhoea, nausea and cough, which were all more frequent in the VR-CAP group (...)*

Similarly, there was a difference in the incidence rates for the following Grade 3 or higher AEs: neutropenia, thrombocytopenia, leukopenia and lymphopenia, which were all more frequent in the VR-CAP group. Of the Grade 3 or higher AEs considered to be related to the study drug, the most common were neutropenia, thrombocytopenia, leukopenia, anaemia, lymphopenia and febrile neutropenia (...)

All SAEs occurred at a frequency of <5% in each group, with the exception of neutropenia, pneumonia and febrile neutropenia. These were also the most commonly reported SAEs considered that were related to any study drug"[2] (Table 5.20).

Table 5.20: LYM-3002 – common AE, safety analysis set
(Table 28 on page 76 of the CS[2])

	VR-CAP (n=240)	R-CHOP (n=242)
Blood and lymphatic system disorders		
Neutropenia, n (%)	211 (87.9)	178 (73.6)
Related to any study drug	209 (87.1)	172 (71.1)
Grade 3 or higher	203 (84.6)	162 (66.9)
Related to any study drug	200 (83.3)	156 (64.5)
Serious	12 (5.0)	13 (5.4)
Related to any study drug	12 (5.0)	12 (5.0)
Thrombocytopenia, n (%)	173 (72.1)	46 (19.0)
Related to any study drug	172 (71.7)	42 (17.4)
Grade 3 or higher	136 (56.7)	14 (5.8)
Related to any study drug	135 (56.3)	12 (5.0)
Serious	8 (3.3)	1 (0.4)
Related to any study drug	8 (3.3)	1 (0.4)
Leukopenia, n (%)	120 (50.0)	93 (38.4)
Related to any study drug	116 (48.3)	87 (36.0)
Grade 3 or higher	105 (43.8)	71 (29.3)
Related to any study drug	103 (42.9)	66 (27.3)
Serious	6 (2.5)	3 (1.2)
Related to any study drug	6 (2.5)	2 (0.8)
Anaemia, n (%)	122 (50.8)	90 (37.2)
Related to any study drug	106 (44.2)	71 (29.3)
Grade 3 or higher	37 (15.4)	33 (13.6)
Related to any study drug	31 (12.9)	27 (11.2)
Serious	4 (1.7)	5 (2.1)
Related to any study drug	4 (1.7)	4 (1.7)
Lymphopenia, n (%)	74 (30.8)	32 (13.2)
Related to any study drug	68 (28.3)	28 (11.6)
Grade 3 or higher	67 (29.7)	21 (8.7)
Related to any study drug	61 (25.4)	17 (7.0)
Serious	1 (0.4)	0
Related to any study drug	0	0
Febrile neutropenia, n (%)	41 (17.1)	34 (14.0)

	VR-CAP (n=240)	R-CHOP (n=242)
Related to any study drug	41 (17.1)	33 (13.6)
Grade 3 or higher	36 (15.0)	33 (13.6)
Related to any study drug	36 (15.0)	32 (13.2)
Serious	26 (10.8)	20 (8.3)
Related to any study drug	26 (10.8)	19 (7.9)
Infections and infestations		
Pneumonia, n (%)	28 (11.7)	15 (6.2)
Related to any study drug	20 (8.3)	11 (4.5)
Grade 3 or higher	17 (7.1)	11 (4.5)
Related to any study drug	13 (5.4)	8 (3.3)
Serious	19 (7.9)	7 (2.9)
Related to any study drug	14 (5.8)	5 (2.1)
General disorders and administration site conditions		
Pyrexia, n (%)	70 (29.2)	37 (15.3)
Related to any study drug	48 (20.0)	23 (9.5)
Grade 3 or higher	8 (3.3)	5 (2.1)
Related to any study drug	7 (2.9)	5 (2.1)
Serious	10 (4.2)	4 (1.7)
Related to any study drug	7 (2.9)	3 (1.2)
Fatigue, n (%)	56 (23.3)	47 (19.4)
Related to any study drug	43 (17.9)	38 (15.7)
Grade 3 or higher	15 (6.3)	6 (2.5)
Related to any study drug	12 (5.0)	5 (2.1)
Serious	0	3 (1.2)
Related to any study drug	0	3 (1.2)
Asthenia, n (%)	38 (15.8)	26 (10.7)
Related to any study drug	29 (12.1)	18 (7.4)
Grade 3 or higher	7 (2.9)	2 (0.8)
Related to any study drug	5 (2.1)	1 (0.4)
Serious	1 (0.4)	0
Related to any study drug	1 (0.4)	0
Oedema peripheral, n (%)	37 (15.4)	25 (10.3)
Related to any study drug	16 (6.7)	13 (5.4)
Grade 3 or higher	1 (0.4)	1 (0.4)
Related to any study drug	1 (0.4)	0
Serious	0	1 (0.4)
Related to any study drug	0	1 (0.4)
Gastrointestinal disorders		
Diarrhoea, n (%)	73 (30.4)	22 (9.1)
Related to any study drug	59 (24.6)	11 (4.5)
Grade 3 or higher	12 (5.0)	5 (2.1)
Related to any study drug	11 (4.6)	4 (1.7)
Serious	4 (1.7)	3 (1.2)
Related to any study drug	4 (1.7)	2 (0.8)
Constipation, n (%)	60 (25.0)	38 (15.7)
Related to any study drug	42 (17.5)	22 (9.1)
Grade 3 or higher	1 (0.4)	2 (0.8)
Related to any study drug	1 (0.4)	1 (0.8)
Serious	0	1 (0.4)
Related to any study drug	0	1 (0.4)
Nausea, n (%)	59 (24.6)	33 (13.6)
Related to any study drug	54 (22.5)	28 (11.6)

	VR-CAP (n=240)	R-CHOP (n=242)
Grade 3 or higher	1 (0.4)	0
Related to any study drug	1 (0.4)	0
Serious	0	0
Related to any study drug	0	0
Vomiting, n (%)	30 (12.5)	13 (5.4)
Related to any study drug	24 (10.0)	8 (3.3)
Grade 3 or higher	1 (0.4)	0
Related to any study drug	1 (0.4)	0
Serious	2 (0.8)	1 (0.4)
Related to any study drug	2 (0.8)	0
Stomatitis, n (%)	26 (10.8)	21 (8.7)
Related to any study drug	20 (8.3)	19 (7.9)
Grade 3 or higher	3 (1.3)	1 (0.4)
Related to any study drug	2 (0.8)	1 (0.4)
Serious	1 (0.4)	2 (0.8)
Related to any study drug	0	2 (0.8)
Nervous system disorders		
Peripheral sensory neuropathy, n (%)	54 (22.5)	48 (19.8)
Related to any study drug	53 (22.1)	45 (18.6)
Grade 3 or higher	12 (5.0)	6 (2.5)
Related to any study drug	12 (5.0)	6 (2.5)
Serious	1 (0.4)	0
Related to any study drug	1 (0.4)	0
Neuralgia, n (%)	25 (10.4)	2 (0.8)
Related to any study drug	25 (10.4)	1 (0.4)
Grade 3 or higher	9 (3.8)	0
Related to any study drug	9 (3.8)	0
Serious	2 (0.8)	0
Related to any study drug	2 (0.8)	0
Respiratory, thoracic and mediastinal disorders, n (%)		
Cough, n (%)	49 (20.4)	20 (8.3)
Related to any study drug	11 (4.6)	6 (2.5)
Grade 3 or higher	3 (1.3)	0
Related to any study drug	2 (0.8)	0
Serious	0	0
Related to any study drug	0	0
Metabolism and nutrition disorders, n (%)		
Decreased appetite, n (%)	46 (19.2)	23 (9.5)
Related to any study drug	36 (15.0)	15 (6.2)
Grade 3 or higher	2 (0.8)	2 (0.8)
Related to any study drug	1 (0.8)	1 (0.4)
Serious	1 (0.4)	0
Related to any study drug	1 (0.4)	0
Skin and subcutaneous tissue disorders		
Alopecia, n (%)	33 (13.8)	33 (13.6)
Related to any study drug	31 (12.9)	33 (13.6)
Grade 3 or higher	2 (0.8)	4 (1.7)
Related to any study drug	2 (0.8)	4 (1.7)
Serious	0	0
Related to any study drug	0	0

	VR-CAP (n=240)	R-CHOP (n=242)
Psychiatric disorders, n (%)		
Insomnia, n (%)	27 (11.3)	18 (7.4)
Related to any study drug	16 (6.7)	8 (3.3)
Grade 3 or higher	1 (0.4)	0
Related to any study drug	1 (0.4)	0
Serious	0	0
Related to any study drug	0	0

Key: CSR, clinical study report; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisone; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisone.
Source: LYM-3002 CSR[21]; Robak et al. 2015.[9]

ERG comment: Results were reported for a safety set (defined as all randomised subjects who received at least one dose of study medication), consisting of 482 (240 of 243 VR-CAP, 242 of 244 R-CHOP). As correctly noted in the CS, "*both chemotherapy induction regimens were generally well tolerated, with low rates of discontinuation due to an AE and low fatality on treatment in both groups*". However, as shown in Table 5.18, more serious AEs were observed for VR-CAP (37.5%) compared to R-CHOP (29.8%) and the serious AEs are usually of higher severity in VR-CAP. While more treatment discontinuations related to the study drug were reported for VR-CAP (7.9%) compared to R-CHOP (5.8%), more fatalities related to R-CHOP (3.0%) compared to VR-CAP (2.0%) were reported. A similar picture is seen in European participants (Table 5.19). It should be noted that more comprehensive results will be available with the final analysis in 2017.

The ERG calculated the risk ratios associated with number of adverse events reported in Table 5.19. Only one of the reported adverse effects reached statistical significance, i.e. when considering only participants from Europe and North America (Canada, USA) significantly more serious adverse events were seen for VR-CAP compared to R-CHOP (Table 5.21). As indicated by the wide confidence interval, only relatively few events were reported (17/68 in R-CHOP group, 33/65 in VR-CAP arm).

Table 5.21: Risk ratios of adverse events

Outcome	Overall publication	Europe and North America
Any TEAE	RR 1.01, 95% CI 0.99 to 1.03	RR 1.01, 95% CI 0.97 to 1.06
Any SAE	RR 1.26, 95% CI 0.98 to 1.62	RR 2.20, 95% CI 1.36 to 3.56
TD AE	RR 1.25, 95% CI 0.67 to 2.30	RR 2.14, 95% CI 0.67 to 6.80
Deaths due to AE	RR 0.95, 95% CI 0.49 to 1.83	N/A

AE = adverse event; CI = confidence interval; N/A = not applicable; RR = risk ratio; SAE = any serious AE; TD = treatment discontinuation due to AE; TEAE = any treatment-emergent AE

Health-related quality of life

In LYM-3002, this endpoint was not assessed as primary or secondary outcome but included as exploratory analysis: "*During the treatment period, no statistically significant or clinically meaningful changes (defined as ≥10 point change*[CS reference 67]*) in EORTC QLQ-C30 global health status were observed within either treatment group or between treatment groups from baseline to cycle 6 or end of treatment*".[21]

ERG comment: No results were reported for the subgroup of participants from Europe and North America.

5.2.3 Included non-RCTs

Eleven non-randomised or non-controlled trials (summarised in Appendix 10 of the CS[2]) were included in the review. According to page 71 of the CS:

"Eight of the identified non-RCTs investigated R-CHOP induction therapy, and were not considered to provide additional relevant evidence further to the pivotal RCT. Three provided evidence for alternative rituximab-chemotherapy induction regimens. Two of these studies investigated R-bendamustine ± cytarabine induction therapy but enrolled too few patients (≤20 previously untreated MCL patients) to provide conclusive supportive evidence and are thus not discussed further.[CS references 77+78] *The remaining study was a large-scale observational study that provided survival data for a number of chemotherapy induction regimens ± rituximab."*[CS reference 79]

ERG comment: None of the non-randomised or non-controlled trials identified data for VR-CAP. By definition none of the trials were randomised or controlled and therefore it would have been inappropriate to use these in any indirect analysis, whereby according to the criteria of Song it is only acceptable to combine and analyse similar trials.[38] Since the primary trial of interest is LYM-3002 and this is a randomised trial, only similar study designs should be considered for indirect analysis.

5.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

Seven included studies were randomised controlled trails (summarised in Tables 5.2 and 5.3) and were eligible for network analysis. As discussed in Section 5.2.4, the non-randomised trials were not eligible for inclusion in a network analysis.

In addition to the LYM-3002 trial, six RCTs met the eligibility criteria of the review and are listed in Table 23 of the CS.[2] According to page 66 of the CS "*Only four RCTs in addition to the LYM-3002 trial reported PFS and/or OS data that could be used in indirect comparison analyses of relevance to the subsequent economic modelling*." The five trials are placed in the network shown in Figure 9 of the CS. According to page 67 of the CS only "*the European MCL Elderly trial*[25] *and the StiL NHL1 trial*[29] *provided evidence of comparators that were included in the decision problem."*

According to page 69 of the CS *"the assumption of comparable patient populations and trial characteristics ...do not hold. Therefore, indirect comparison outcomes should be interpreted with caution*." In addition the CS did not perform network analysis because it was not

deemed appropriate given the lack of similarity in the patient populations and trial characteristics (discussed pages 67-69 CS).

ERG comment: The ERG agrees with the CS that only three trials (LYM-3002, European MCL Elderly trial and StiL NHL) were eligible for indirect comparison and network analysis according to the decision problem. The following randomised controlled trials were excluded from the network for the following reasons:

- GLSG, LY05 and NCRI III were excluded because they had comparators irrelevant to the final scope (CHOP and FC).
- The BRIGHT study[CS reference 72] was not included because it included mantle cell lymphoma patients and indolent non-Hodgkin's lymphoma patients but did not report separate data for mantle cell lymphoma only.

The three included trials should only be combined in an indirect comparison or network analysis if they are considered similar in terms of population, posology (for common comparators) and outcome. Therefore the ERG further critiqued the similarity of the three trials, summarised in Table 5.22. The table illustrates that whilst the patient characteristics and quality of LYM-3002 and the European MCL network can be considered quite similar the design of the European MCL network from induction to maintenance phase compromises the statistical analysis of the survival outcomes. Whilst overall the StiL NHL trial is also similar to LYM-3002, it is compromised by the lack of statistical powering for the MCL subpopulation and lack of clear reporting for the characteristics of these patients.

Table 5.22: Brief summary of similarity of trials included in network meta-analysis

	LYM-3002[9]	**European MCL network**[25] **Induction phase**	**StiL NHL**[29]
Population	Aged 26-88 MCL Stage II, III or IV; ineligible for HSCT (but 15.6-17.2% were eligible); ECOG ≤ 2 N= 487	Aged 60-87 years Newly diagnosed MCL (Ann Arbor stage II to IV); ineligible for HSCT; ECOG ≤ 2 N= 485	Aged 31-83 Newly diagnosed indolent lymphoma or MCL (stage II to IV). MCL 18-19% of total population HSCT eligibility not reported ECOG not reported N= 95 for MCL
Posology	VR-CAP for up to 8 cycles Or R-CHOP for up to 8 cycles	Randomised to 6 cycles of R-FC induction therapy or 8 cycles of R-CHOP induction therapy Patients responding to induction treatment were randomised to maintenance treatment: interferon alpha or ritiuximab	R-bendamustine for up to 6 cycles. Or R-CHOP for up to 6 cycles
Quality Limitations	High risk of bias for blinding and moderate	High risk of bias for blinding and moderate	Unclear reporting of characteristics and

	LYM-3002[9]	European MCL network[25] Induction phase	StiL NHL[29]
	risk for balance of baseline characteristics	risk for balance of baseline characteristics	analysis for MCL subpopulation. High risk of bias for blinding and for statistical power calculation
Outcomes eligible (Follow-up)	See table 5.2	Time to Treatment Failure (84 months) Overall Survival (84 months) Safety (unclear)	Progression free survival (96 months) Overall response rate Complete response Partial response
Limitations noted by CS		Not designed to assess the clinical efficacy of induction therapy ± maintenance therapy, survival estimates cannot be fully disaggregated and should be interpreted with caution. PD, assessments were only conducted twice yearly during follow-up in the European MCL Elderly trial compared with every 6-8 weeks in LYM-3002 (Page 68 CS)	PD assessments were only conducted twice yearly during follow-up in the StiL NHL1 trial.
Key: CS = company's submission; ECOG = Eastern Cooperative Oncology Group; HSCT = haematopoietic stem cell transplantation; MCL = mantle cell lymphoma; NHL = non-Hodgkin's lymphoma; PD = progressive disease; R-bendamustine = Rituximab with bendamustine; R-CHOP = Rituximab, cyclo-phosphamide, doxorubicin, vincristine and prednisolone; VR-CAP = Bortezomib (Velcade®) in combination with rituximab, cyclo-phosphamide, doxorubicin and prednisone			

The ERG agrees with the CS that the results of the indirect analysis should be treated with caution due to the lack of similarity between the three included trials, and that they are not similar enough to justify performing a network analysis.

5.4 Critique of the indirect comparison and/or multiple treatment comparison

While no multiple treatment comparison was conducted, results of several indirect comparisons were presented as part of the CS[2] and in response to the clarification letter[16]. The company emphasises shortcomings of the indirect comparisons: "*It is important to note that established practice for the treatment of patients with previously untreated MCL for whom HSCT is unsuitable in NHS England and Wales is the use of R-CHOP induction therapy. Other induction therapy regimens listed in the final appraisal scope are not considered to be relevant comparators for VR-CAP as these are generally reserved for patients who cannot tolerate R-CHOP and, therefore, VR-CAP (...) Nevertheless, to align*

with the final appraisal scope, and for transparency, indirect comparison analyses to alternative rituximab-chemotherapy induction regimens have been conducted where possible. These analyses are not considered robust due to important heterogeneity between LYM-3002 and the comparator studies, and due to methodological limitations of the comparator studies".[2]

In the CS, the company presented the results of two indirect comparisons using the Bucher method for progression-free and overall survival.[39] In all comparisons a link between VR-CAP and R-CHOP was provided by the LYM-3002 trial.[9] A link between R-bendamustine and R-CHOP, allowing an indirect comparison between VR-CAP and R-bendamustine, was based on the StiL NHL 1 trial which provided data for both, progression-free and overall survival.[29] Another trial, European MCL network, reported results for overall survival for the comparison of R-CHOP and R-FC, allowing an indirect of VR-CAP and R-FC for this outcome[25] (table 5.23).

Table 5.23: Bucher indirect comparison analyses of PFS and OS
(Table 24 on page 70 of the CS[2])

	HR	Lower 95% CI	Upper 95% CI
Overall survival			
VR-CAP vs. R-FC	0.53	0.35	0.81
VR-CAP vs. R-bendamustine	0.63	0.31	1.25
Progression-free survival			
VR-CAP[a] vs. R-bendamustine	1.04	0.59	1.84

Key: CI, confidence interval; HR, hazard ratio; IRC, independent review committee; NHL, non-Hodgkin's lymphoma; OS, overall survival; PFS, progression-free survival; R-bendamustine, rituximab with bendamustine; R-FC, rituximab with fludarabine and cyclophosphamide; StiL, Study Group of indolent Lymphomas; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisone.
Notes: [a], PFS investigator data used in indirect comparison as StiL NHL1 did not use IRC assessment.

In response to the clarification letter, the company provided further indirect comparisons for overall response rate and complete response. However, the company reiterated "the limitations of indirect comparison including these studies described in the company submission, given the between-study heterogeneity and, with respect to R-bendamustine, the small size and lack of detail on patient characteristics for the subgroup of MCL patients within the StiL NHL1 trial"[16] (Table 5.24).

Table 5.24: Bucher indirect comparison analyses of ORR and CR
(Table 4 on page 8 of the response to the clarification letter[16])

	OR	Lower 95% CI	Upper 95% CI
Overall response rate			
VR-CAP vs. R-FC	2.42	1.09	5.39
VR-CAP vs. R-bendamustine	3.54	0.61	20.69
Complete response			
VR-CAP vs. R-FC	1.22	0.73	2.05
VR-CAP vs. R-bendamustine	1.69	0.67	4.28

Key: CI, confidence interval; CR, complete response; OR, odds ratio; ORR, overall response rate; R-bendamustine, rituximab with bendamustine; R-FC, rituximab with fludarabine and cyclophosphamide; StiL, Study Group of indolent Lymphomas; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisone.
Notes: Investigator assessments for R-FC and R-bendamustine vs. Independent Review Committee for VR-CAP

ERG comment: As detailed in Section 5.3, the ERG agrees with the CS that the results of the indirect analysis should be treated with caution due to the lack of similarity between the three included trials. The wide confidence intervals reported above could partly be explained by this heterogeneity. It should also be noted that the three trials included in the indirect analyses are linked to high risk of bias, as detailed in Section 5.1.4.

Other potential comparators that could have been included in a potential network or as part of indirect comparisons are R-CVP and R-chlorambucil. The ERG sought clarification why these were not included and it was stated that "*these are only used to treat frailer patients who would not be eligible for VR-CAP*".[16] According to the advisory board which the company used for the project, the same could be said of R-FC which was a comparator: *"R-FC and R-chlorambucil are only used in very frail, elderly patients"*.[7]

However, as R-CVP and R-chlorambucil were outside the final scope, the decision by the company not to include these comparators is correct.[1]

5.5 Additional work on clinical effectiveness undertaken by the ERG

As the main evidence comes from one randomised controlled trial, LYM-3002, no additional analyses linked to the clinical effectiveness were conducted by the ERG.

The ERG calculated risk ratios for the adverse effects reported in the LYM-3002 trial.

5.6 Conclusions of the clinical effectiveness section

The risk of bias of the only RCT, LYM-3002, assessing VR-CAP (including bortezomib) as well as the two additional studies included in the indirect comparison was judged to be high. Therefore, results should be interpreted with caution.

For LYM-3002 only interim results (December 2013) are available, the final analysis is planned for 2017. Furthermore, this trial did not include any participants from the UK and only approximately one third of the participants were from Europe or North America. Focussing on this subgroup, results for all endpoints will be underpowered.

The majority of searches in the CS were well documented and easily reproducible; searches were carried out in line with the NICE guide to the methods of technology appraisal sections

5.2.2 and 5.2.4. In addition, relevant references have been excluded. However, as these were linked to the LYM-3002 trial, no relevant data were excluded.

The main results for the endpoints defined in the final scope for the direct comparison of VR-CAP with R-CHOP are presented in Table 5.25.

Table 5.25: Summary of interim analysis results for LYM-3002: direct analysis of VR-CAP versus R-CHOP

Outcome from final scope	Overall Population	Western Europe and North America
Overall survival	HR = 0.80, 95% CI 0.59 to 1.10[§]	HR = 1.09; 95% CI: 0.47, 2.52[^]
Progression-free survival	HR = 0.63, 95% CI 0.50 to 0.79[□]	HR = 0.77, 95% CI 0.43 to 1.38[§]
Overall response rate	OR = 1.43, 95% CI 0.75 to 2.72[§]	OR = 1.16, 95% CI 0.32 to 4.25[§]
Duration of response (median duration, days)	VR-CAP: 1110 (95% CI 813 to 1420) R-CHOP: 459 (95% CI 379 to 518)	N/A
Time to new anti-lymphoma treatment/time to progression	HR = 0.58, 95% CI 0.45 to 0.74[$]	HR = 0.63, 95% CI 0.34 to 1.16[§]
AEs of treatment*: **Any TEAE** **Any SAE** **TD AE** **Deaths due to AE**	 RR = 1.01, 95% CI 0.99 to 1.03[^] RR = 1.26, 95% CI 0.98 to 1.62[^] RR = 1.25, 95% CI 0.67 to 2.30[^] RR = 0.95, 95% CI 0.49 to 1.83[§]	 RR[‡] = 1.01, 95% CI 0.97 to 1.06[^] RR[‡] = 2.20, 95% CI 1.36 to 3.56[Δ] RR[‡] = 2.14, 95% CI 0.67 to 6.80[^] N/A
Health-related quality of life (EORTC-QLQ-C30)	No difference[#]	N/A

AE = adverse event; CI = confidence interval; EORTC-QLQ-C30 = European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire; HR = hazards ratio; NA = not assessed; OR = odds ratio; OS = overall survival; R-CHOP = Rituximab, cyclo-phosphamide, doxorubicin, vincristine and prednisolone; RR = risk ratio; SAE = any serious AE; TD = Treatment discontinuation due to AE; TEAE= Any treatment-emergent AE; VR-CAP, bortezomib (Velcade®) with rituximab, cyclophosphamide, doxorubicin and prednisone
[§] In favour of VR-CAP, not statistically significant
[□] In favour of VR-CAP, statistically significant
[^] In favour of R-CHOP, not statistically significant
[Δ] In favour of R-CHOP, statistically significant
* AE risk ratios were calculated by the ERG
[‡] Data calculated from European Union and North America subgroups combined (n=150)
[#] None of the observed changes in scores through the end of treatment were considered statistical significant or clinically meaningful (defined as ≥10 point change)

Three trials were identified for indirect analysis of VR-CAP versus R-FC or VR-CAP versus R-bendamustine. Results are summarised in table 5.26. Outcomes of the final scope are presented for the overall population and for the subgroup of Western Europe and North America.

Table 5.26: Indirect analysis of VR-CAP versus R-FC or VR-CAP versus R-bendamustine

Outcome from final scope	Overall Population	Western Europe and North America
Overall survival: VR-CAP vs. R-FC	HR = 0.53, 95% CI 0.35 to 0.81	N/A

VR-CAP vs. R-bendamustine	HR = 0.63, 95% CI 0.31 to 1.25	N/A
Progression-free survival: VR-CAP vs. R-FC VR-CAP vs. R-bendamustine	 N/A HR = 1.04, 95% CI 0.59 to 1.84	 N/A N/A
Overall response rate VR-CAP vs. R-FC VR-CAP vs. R-bendamustine	 OR = 2.42, 95% CI 1.09 to 5.39 OR =3.54, 95% CI 0.61 to 20.69	 N/A N/A
Duration of response	N/A	N/A
Time to new anti-lymphoma treatment/time to progression	N/A	N/A
Adverse effects of treatment*: **Any TEAE n (%)** **Any SAE n (%)** **TD AE n (%)** **Deaths due to AE n (%)**	N/A	N/A
Health-related quality of life	N/A	N/A
AE = adverse event; CI = confidence interval; HR = hazards ratio; NA = not assessed; OR = odds ratio; OS = overall survival; R-CHOP = R-CHOP = Rituximab, cyclo-phosphamide, doxorubicin, vincristine and prednisolone; R-FC = rituximab with fludarabine and cyclophosphamide; RR = risk ratio; SAE = any serious AE; StiL = Study Group of indolent Lymphomas; TD = Treatment discontinuation due to AE; TEAE= Any treatment-emergent AE; VR-CAP, bortezomib (Velcade®) with rituximab, cyclophosphamide, doxorubicin and prednisone		

6. Assessment of cost-effectiveness

6.1 ERG comment on company's review of cost-effectiveness evidence

6.1.1 Searches for cost effectiveness review

A literature review was conducted to identify any existing economic analyses in this area in order to address the decision problem and inform the economic model. Searches were reported for MEDLINE, MEDLINE In-Process, Embase, NHS EED and EconLit. No date or language limits were applied. Conference abstract searching was carried out as part of the clinical effectiveness work reported in Section 4.1 and Appendix 3 of the CS.[2, 15]

The database hosts for each database, date span and search dates were stated. Update search dates and numbers retrieved were also provided. The searches were on the whole well-reported and reproducible. The database searches were clearly structured and used combinations of index terms and free text. The study design limits applied to identify economic evidence in Medline and Embase were applied appropriately and appeared to be those developed by SIGN.[14]

Measurement and valuation of health effects

A systematic review of health-related quality of life (HRQOL) studies was conducted for mantle cell lymphoma (MCL) and non-Hodgkin's lymphoma (NHL). Searches were reported for MEDLINE, MEDLINE In-Process, Embase, EconLit and NHS EED. As reported in Section 5.1 for the clinical effectiveness searches, five conference proceeding websites were also searched. The host, date span and search dates were reported for all resources. Update search dates and numbers retrieved were also provided. The searches were well reported and reproducible.

The ERG noted several syntax errors in the Cochrane Library searches, where attempts were made to truncate NHL free-text terms using Ovid syntax (lines 3-4, 6-8 on page 144 of the CS).[15] These errors will have impaired recall of relevant NHL references. The MCL component of the search was not negatively impaired.

A facet of HRQOL terms was included in the NHS EED search. Although these terms were comprehensive, the ERG considered this limit to be overly restrictive when applied to small content-specific resources, such as NHS EED. The ERG reproduced the company's NHS EED search, minus the HRQOL terms, and screened the resulting 59 records, 44 of which were missed by the company's original search. The ERG identified relevant references which were missed by the CS search, however these were found to be secondary publications to primary studies already flagged by the company.

Resource identification, measurement and valuation studies

A search to retrieve references to resource allocation and MCL was conducted. Searches were reported for MEDLINE, MEDLINE In-Process, Embase, EconLit and NHS EED. As reported in Section 5.1 for the clinical effectiveness searches, five conference proceeding websites were also searched. The host, date span and search dates were reported for all resources. Update search dates and numbers retrieved were also provided. The searches were well reported and reproducible.

The database searches were clearly structured and used combinations of index terms and free text. The ERG noted that the terms to identify resource allocation were very limited, and it was possible that inclusion of additional terminology and synonyms may increase the recall of the Medline and Embase searches.

A limited facet of resource allocation terms was included in the NHS EED search. The ERG considered this limit to be overly restrictive when applied to small content-specific resources, such as NHS EED. The ERG reproduced the company's NHS EED search, minus the resource allocation terms, and screened the resulting five records, all of which were missed by the company's original search.

6.1.2 Inclusion/exclusion criteria used in the study selection

Table 6.1 presents the inclusion criteria used for the review.

Table 6. 1: Inclusion criteria for cost-effectiveness studies

Types of studies	Studies reporting cost-effectiveness analyses and results were not filtered by study design. All studies were included barring reviews, letters and comment articles.
Types of participants	Previously untreated adult MCL patients.
Types of intervention	First-line treatment; any intervention with active treatment could be included to allow the identification of additional potentially relevant models.
Types of outcomes	Incremental costs and QALYs or any other natural unit measure of effectiveness (e.g. life years) reported together with costs.
Key: MCL, mantle cell lymphoma; QALY, quality-adjusted life year.	

ERG comment: These criteria seem to be appropriate for the purpose of this review.

6.1.3 Included/excluded studies in the cost-effectiveness review

In total 234 potentially relevant studies were identified of which zero remained after exclusion of duplicates (59) and reviewing title and abstracts.

6.1.4 Conclusions of the cost-effectiveness review

There were no relevant studies identified in the literature that assess the cost-effectiveness of treatments for MCL.

ERG comment: The ERG agrees with the conclusions from the company that none of the selected studies were relevant for the decision problem.

6.2 Summary and critique of company's submitted economic evaluation by the ERG

The ERG has assessed the company's economic evaluation using the Philips et al checklist for quality assessing decision analytic models.[40] This is shown in Appendix 2 of this report and is used to assist the narrative critique in the following sections.

An overall summary of the *de novo* economic model developed by the company is given in Table 6.2.

Table 6.2: Summary of the company's economic evaluation (with signposts to CS)

	Approach	Source / Justification	Signpost (location in CS)
Model	Markov model with transition probability based upon survival functions of OS, PFS and TFI. Remaining transition probabilities are derived from the average duration of PFS after second-line treatment.	This modelling approach is commonly used for the cost-effectiveness analysis of treatment in other cancer types (including lymphoma). Furthermore, a Markov model is especially suitable for diseases in which patients progress through distinct and definable stages of disease. MCL is such a disease type.	Section 5.2, page 94
States and events	Progression-free survival from first line treatment Progressed from first-line treatment Progression-free survival from second-line treatment Progressed from second-line treatment Death	The company considers progression to be the best surrogate for OS and a key predictor for changes in resource use. Additional treatment lines (after 2 lines) were excluded as it will increase complexity and uncertainty. Furthermore, the impact is considered marginal as treatment is palliative	Section 5.2, page 94-95
Comparators	R-CHOP, R-FC, R-bendamustine	R-CHOP is the established standard of care in patients with newly diagnosed MCL for whom SCT is unsuitable. R-FC and R-bendamustine are provided if patients are not fit enough to receive R-CHOP or are contraindicated.	Section 5.2, page 96
Natural History	Derived from estimates regarding PFS and OS. OS is differentiated into survival for patients with progression (PPS) and without progression (PrePS). Consequently, survival is initially estimated from the PrePS curve until patients experience progression. In addition	The company argued that the fit of any survival curve to OS data from the LYM-3002 trial[9] was poor and that therefore it was better to fit curves separately for PrePS	Section 5.3, page 102-113

	Approach	Source / Justification	Signpost (location in CS)
	to PrePS, background mortality is added to include death due to other reasons than MCL After progression, survival is derived from the PPS curve. As PFS includes both progression and death as events, the difference between PFS and PrePS is used to determine whether patients progress.	and PPS states.	
Treatment effectiveness	Treatment influences overall survival for patients without progression, progression-free survival and treatment-free interval.	Differences in the end-points were observed between R-CHOP and VR-CAP in the LYM-3002 trial.[9] The company argued that there was no difference in OS during the PPS state and so the treatment effect, estimated from the LYM-3002 trial, was only applied to OS in the PrePS state.	Section 5.3. page 103, 104, 109, 114
Adverse events	The total number of AEs were divided by the total years on treatment (= average number of weeks on treatment x total number of patients) to estimate the yearly incidence of AEs. This yearly incidence is translated to a cycle-based probability. The impact of AE is included in both the costs and quality of life during first-line treatment.	Based on observed AEs of at least grade 3 with occurrence in more than 5% of the population in the LYM-3002 trial for the R-CHOP and VR-CAP arm.[9] Incidence of AE for R-FC and R-bendamustine are derived from Kluin-Nelemans et al.[25] and Flinn et al.,[24] respectively.	Section 5.4, page 134-136 and Section 5.5, page 157
Health related QoL	Utility values were assigned to the four 'living' health states	Utility values for PFS from 1st line treatment and progressed after 1st line treatment were estimated from the quality of life measured in the LYM-3002 trial. Utility value for PFS from 2nd line treatment was considered similar to PFS from 1st line treatment. Utility value for progression after	Section 5.4, page 122-124 (Table 42), 137-138 (Table 47)

	Approach	Source / Justification	Signpost (location in CS)
		2nd line treatment is derived from Doorduijn et al.[41]	
Resource utilisation and costs	Costs consisted of four categories: drug acquisition and administration costs, health state costs, adverse events costs and end-of life costs. Drug acquisition costs were adjusted for the relative dose intensity of the different treatments and wastage was included (assuming no vial sharing). Health state costs include the costs of tests, outpatient visits, concomitant medication and transfusions	Based on UK reference costs, literature and expert opinion	Section 5.5, page 140-158
Discount rates	A 3.5% discount rate was used for both costs and effects.	According to NICE reference case	Section 5.2, page 96
Sub groups	No subgroup analysis has been performed	No subgroups were identified where the effectiveness of VR-CAP was significantly different from the ITT analysis	Section 5.9, page 182
Sensitivity analysis	One-way deterministic sensitivity analysis, scenario analyses and probabilistic sensitivity analysis	Ranges based on confidence intervals and assumptions.	Values and distribution input parameters: Appendix 21. Results section 5.8, page 170
Key: AE, adverse event; ITT, intention to treat; OS, overall survival; PFS, progression-free survival; PPS, post-progression survival; PrePS, pre-progression survival; Qol, quality of life; R-bendamustine, rituximab with bendamustine; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisone; R-FC, rituximab with fludarabine and cyclophosphamide; SCT, stem cell transplant; TFI, treatment-free interval; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisone			

6.2.1 NICE reference case checklist (TABLE ONLY)

Table 6.3: NICE reference case checklist

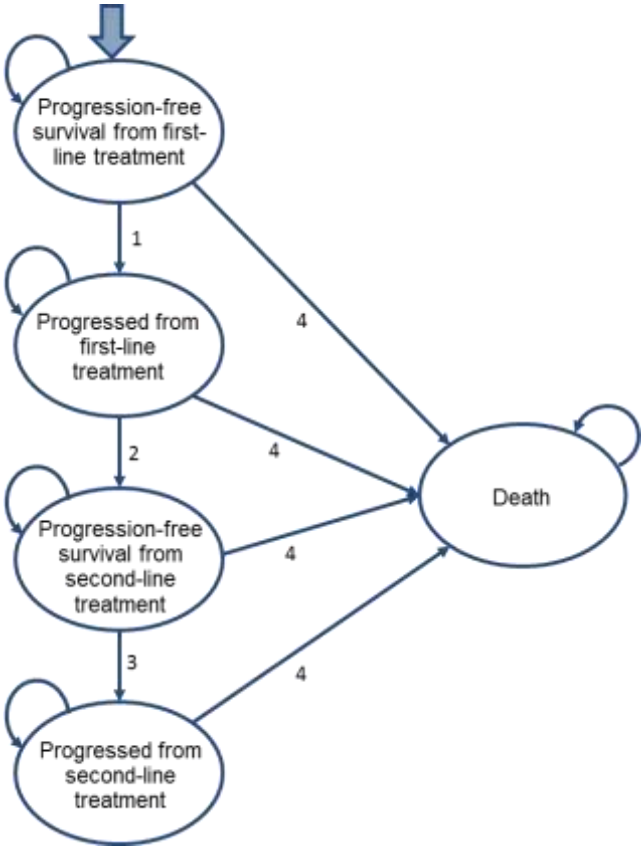
Elements of the economic evaluation	Reference Case	Included in submission	Comment on whether *de novo* evaluation meets requirements of NICE reference case
Population	The NICE scope defined: *Adults with* previously untreated mantle cell lymphoma, who are unsuitable for haematopoietic stem cell transplantation.	Yes	

Comparator(s)	Therapies routinely used in the NHS, including technologies regarded as current best practice	Partially	R-FC and R-bendamustine were included as comparators in an exploratory analysis as these treatments are often administered to patients not deemed fit enough, or contraindicated, to receive CHOP-like therapy
Type of economic evaluation	Cost-effectiveness analysis	Yes	
Perspective on costs	NHS and PSS	Partially, see comment	End-of-life costs include costs for 'third sector' (i.e. not-for-profit and non-governmental) healthcare organisations
Perspective on outcomes	All health effects on individuals	Yes	
Time horizon	Sufficient to capture differences in costs and outcomes	Yes	Time horizon is 20 years.
Synthesis of evidence in outcomes	Systematic review	No	No synthesis was performed for any of the input parameters
Measure of health effects	QALYs	Yes	
Source of data for measurement HRQoL	Reported directly by patients and/or carers	Yes	Based upon the answers on the clarification letter.
Source of preference data for valuation of changes in HRQoL	Sample of public	Yes	
Discount rate	Annual rate of 3.5 on costs and health effects	Yes	
Equity weighting	No special weighting	Yes	
Sensitivity analysis	Probabilistic sensitivity analysis	Partially	Not all potentially stochastic parameters were incorporated as stochastic (see section 6.2.10)
HRQoL= Health-related Quality of Life; NHS= National Health Services; NICE= National Institute for Health and Care Excellence; PSS= Personal Social Services; QALY= Quality-adjusted life year			

6.2.2 Model structure

A de novo cost-effectiveness model was developed as a previously developed model was not available in the literature. The model has a Markov health state structure and consists of five health states: progression-free survival (PFS) from first-line treatment, progression from first-line treatment, PFS from second-line treatment, progression from second-line treatment and death. A graphical representation is shown in Figure 6.1.

Figure 6.1: Model diagram



Key: OS, overall survival; PFS, progression-free survival; PPS, post-progression survival; PrePS, pre-progression survival;
Notes: 1. Modelled using curve fit to PFS Kaplan–Meier data; 2. Modelled using curve fit to treatment-free interval; 3. Modelled using average PFS from second-line treatment; 4. Modelled using curve fit to OS Kaplan–Meier curve (PPS and PrePS plus general population background mortality data).More information can be found in section 6.2.6.

All patients start in the health state 'PFS from first-line treatment'. Subsequently, patients may experience different phases of the disease from 'progression from first-line' to 'PFS from second-line treatment' and 'progression from second-line treatment'. The model distinguishes 'progressed from first-line treatment' from 'PFS from second-line treatment' as second-line treatment does not start immediately after disease progression. The progression-free health states (both from first-line and second-line treatment) distinguish patients on and off active treatment.

Patients can die from all 'alive' health state, with a different probability of death from 'PFS from first-line treatment' as compared to the probability of death from other health states. All transition probabilities were estimated from data of the LYM-3002 trial. Detailed information about the transition probabilities between health states can be found in Section 6.2.6.

Any additional treatment lines are not included in the model as the inclusion would increase complexity and uncertainty. Furthermore, only a minority of the patients require a third treatment line (20.3% in the LYM-3002 trial) and only palliative treatment will be administered at that time.

A cycle length of one week was used in the model as it was considered sufficient to capture short-term changes in progression status. No half-cycle correction was applied in the model. A motivation for the exclusion was not provided in the company's submission.

ERG comment: The newly developed model follows a logical structural with respect to the nature of the disease. The ERG identified only two possible concerns: the exclusion of the half-cycle correction and the exclusion of any additional treatment lines after second-line treatment. After the request of the ERG in the clarification letter, the company motivated that the half-cycle correction was excluded as it would only have a marginal impact on the incremental cost-effectiveness ratio (ICER) due to the short cycle length (one week). The company provided the results of the inclusion of the half-cycle correction in the model and the ICER increased from £20,362 to £21,293. However, after comparing the results of the model with and without half-cycle correction, the ERG is not convinced that the half-cycle correction was implemented correctly. Therefore the ERG has made its own correction for the new ERG base case (see sections 6.3 and 7). The exclusion of any additional treatment lines seems reasonable considering the lack of evidence of treatment efficacy and the minority of patients receiving a third treatment line. Furthermore, the ERG considers the exclusion of any additional treatment line as conservative, because it is expected that a third treatment line is more often administered to the comparators (R-CHOP, R-FC, R-bendamustine) due to a worse PFS after first-line treatment and similar effectiveness of second-line treatment. In addition, it is assumed that the clinical effect of any third line treatment is already incorporated in the observed OS. Consequently, the inclusions of a third line treatment might decrease the incremental costs and therefore the ICER.

6.2.3 Population

The patient population in the model was in line with the licensed population for bortezomib: "adults with previously untreated mantle cell lymphoma, who are unsuitable for haematopoietic stem cell transplantation"[2] The licensed indication was granted by the European Medicines Agency (EMA): 'Bortezomib in combination with rituximab, cyclophosphamide, doxorubicin and prednisone (VR-CAP) is indicated for the treatment of adult patients with previously untreated mantle cell lymphoma (MCL), an incurable subtype of non-Hodgkin's lymphoma, who are unsuitable for haematopoietic stem cell transplantation (HSCT)'[42, 43].

ERG comment: As detailed in Section 5.2.1 of this report, there is uncertainty regarding the generalisability of the trial population to clinical practice in the UK.

6.2.4 Interventions and comparators

The intervention described in the CS ('bortezomib') matches the intervention described in the final scope. The EMA Summary of Product Characteristics (SmPC)[44] provides the following details regarding administering the intervention: "*1 mg powder for solution for injection is administered intravenously at the recommended dose of 1.3 mg/m^2 body surface area twice weekly for two weeks on days 1, 4, 8, and 11, followed by a 10-day rest period on days 12-21. This 3-week period is considered a treatment cycle. Six cycles are recommended, although for patients with a response first documented at cycle 6, two additional cycles may be given. At least 72 hours should elapse between consecutive doses of bortezomib*.

The following medicinal products are administered on day 1 of each bortezomib 3-week treatment cycle as intravenous infusions: rituximab at 375 mg/m^2, cyclophosphamide at 750 mg/m^2 and doxorubicin at 50 mg/m^2. Prednisone is administered orally at 100 mg/m^2 on days 1, 2, 3, 4 and 5 of each bortezomib treatment cycle.

Prior to initiating a new cycle of therapy:

- *Platelet counts should be ≥ 100,000 cells/µL and the absolute neutrophils count (ANC) should be ≥ 1,500 cells/µL*
- *Platelet counts should be ≥ 75,000 cells/µL in patients with bone marrow infiltration or splenic sequestration*
- *Haemoglobin ≥ 8 g/dL*
- *Non-haematological toxicities should have resolved to Grade 1 or baseline.*

Bortezomib treatment must be withheld at the onset of any ≥ Grade 3 treatment-related non-haematological toxicities (excluding neuropathy) or ≥ Grade 3 haematological toxicities." Dose adjustments are presented in Table 5 (page 7) of the SmPC.[44]

Granulocyte colony stimulating factors may be administered for haematological toxicity according to local standard practice. Prophylactic use of granulocyte colony stimulating factors is considered in case of repeated delays in cycle administration. Platelet transfusion for the treatment of thrombocytopenia is considered when clinically appropriate.

Bortezomib demonstrates a synergistic and additive effect when administered with other clinically active agents in MCL. The established standard of care in the National Health Service (NHS) England and Wales is rituximab with cyclophosphamide, doxorubicin, vincristine and prednisone (R-CHOP). Other comparators included in the scope are bendamustine plus rituximab (with or without cytarabine) and rituximab, fludarabine and cyclophosphamide (R-FC). These comparators are indicated for patients with frailty due to advanced age and/or comorbidities. Thus, R-bendamustine and R-FC are not used in the same placement as VR-CAP. In the model, these comparators were not included in the base-case, presented as 'exploratory' only.

R-maintenance is used in clinical practice in patients with response to induction therapy. No maintenance treatment is assumed in the model base case. However, the potential impact of introducing R-maintenance followed by induction therapy with VR-CAP versus R-CHOP was explored. Several scenario and threshold analyses were performed to assess the potential impact of R-maintenance on the cost effectiveness of VR-CAP induction therapy.

ERG comment: The exclusion of R-FC and R- bendamustine in the base case is inconsistent with the scope. Therefore, the ERG has included these comparators in all analyses, not only as 'exploratory comparators'.

Although the effectiveness of R-maintenance therapy is difficult to assess, the exclusion of post-induction R-maintenance treatment by the company is debatable. During the scoping meeting it was suggested by clinical experts that post induction R-maintenance treatment is common practice in the UK. Thus, the ERG considers it relevant to explore the impact of R-

maintenance treatment by applying the scenario analyses presented by the company to the ERG base case.

6.2.5 Perspective, time horizon and discounting

A lifetime horizon of 20 years is used in the base-case analysis. Costs are considered from the NHS and PSS perspective. Both costs and effects are discounted at 3.5%. The impact of the time horizon was assessed in scenario analyses with a time horizon of 10, 15, 25 and 30 years.

ERG comment: The ERG concludes that the discount rate and perspective are in line with the NICE reference case. Considering the average age of 69 years in the LYM-3002 trial[9] and that the median survival is less than five years a time horizon of 20 years is considered adequate and similar to a lifetime perspective. Nevertheless, a longer time horizon is feasible as not all patients died within 20 years. A longer time horizon will further decrease the ICER. Therefore, the choice for 20 years instead of a longer time horizon is conservative.

6.2.6 Treatment effectiveness and extrapolation

All data regarding treatment effectiveness is derived from the LYM-3002 trial.

In principal, the area-under-the curve (AUC) method was used to estimate transition probabilities between the health states, with the exception of the transition from 'progression-free survival from second-line treatment' to 'progressed from second-line treatment'. That transition probability was derived from the average duration of PFS from second-line treatment. Table 6.4 summarises how the different transition probabilities are estimated. More information about each estimation is found below.

Table 6.4: Transition probabilities in the Markov model

Transition probability		Nr	Method of calculation
From	**To**		
PFS from 1st line	PFS from 1st line		PFS
PFS from 1st line	Death	4	OS in patients without progression (PrePS) + background mortality
PFS from1st line	Progression 1st line	1	PFS – (PrePS + background mortality)
Progression 1st line	Death	4	OS in patients with progression (PPS)
Progression 1st line	PFS from 2nd line	2	Treatment-free interval (TFI)
Progression 1st line	Progression 1st line		1 – TFI - PPS
PFS from 2nd line	Progression 2nd line	3	Constant transition probability derived from the average duration of PFS in 2nd line p_prog=exp(-1/duration PFS in weeks)
PFS from 2nd line	Death	4	PPS
PFS from 2nd line	PFS from 2nd line		1 – PPS – p_prog
Progression 2nd line	Death	4	PPS
Progression 2nd line	Progression 2nd line		1 - PPS

Key: PFS = progression-free survival, PPS = post-progression survival (=OS in patients with progression); PrePS = pre-progression survival (=OS in patients without progression); p_prog = probability progression 2nd line; OS = overall survival; TFI = treatment-free interval

Patients in both PFS health states (from first and second line treatment) are distinguished in two groups: patients on and off treatment. The probability of going off treatment from the first treatment line is derived from the proportion of patients receiving different treatment

cycles as observed in the LYM-3002 trial.[9] The probability of going off treatment from the second treatment line is derived from the average duration of second line treatment (90 days).

With respect to the AUC method, three main survival functions were estimated: 1) progression-free survival (PFS), 2) overall survival (OS) and 3) treatment-free interval (TFI). Detailed information about the definition of the end-points is reported in table 6.5. Different parametric survival functions (exponential, Weibull, log-normal, log-logistic, gamma and Gompertz distribution) were estimated for all outcomes measures. The selection of the most appropriate distribution was based upon statistical fit (AIC/BIC), visual fit and long-term fit. More specific information about the extrapolation of all these outcomes measures are presented in the subsequent paragraphs.

Table 6.5: Definition of end-points

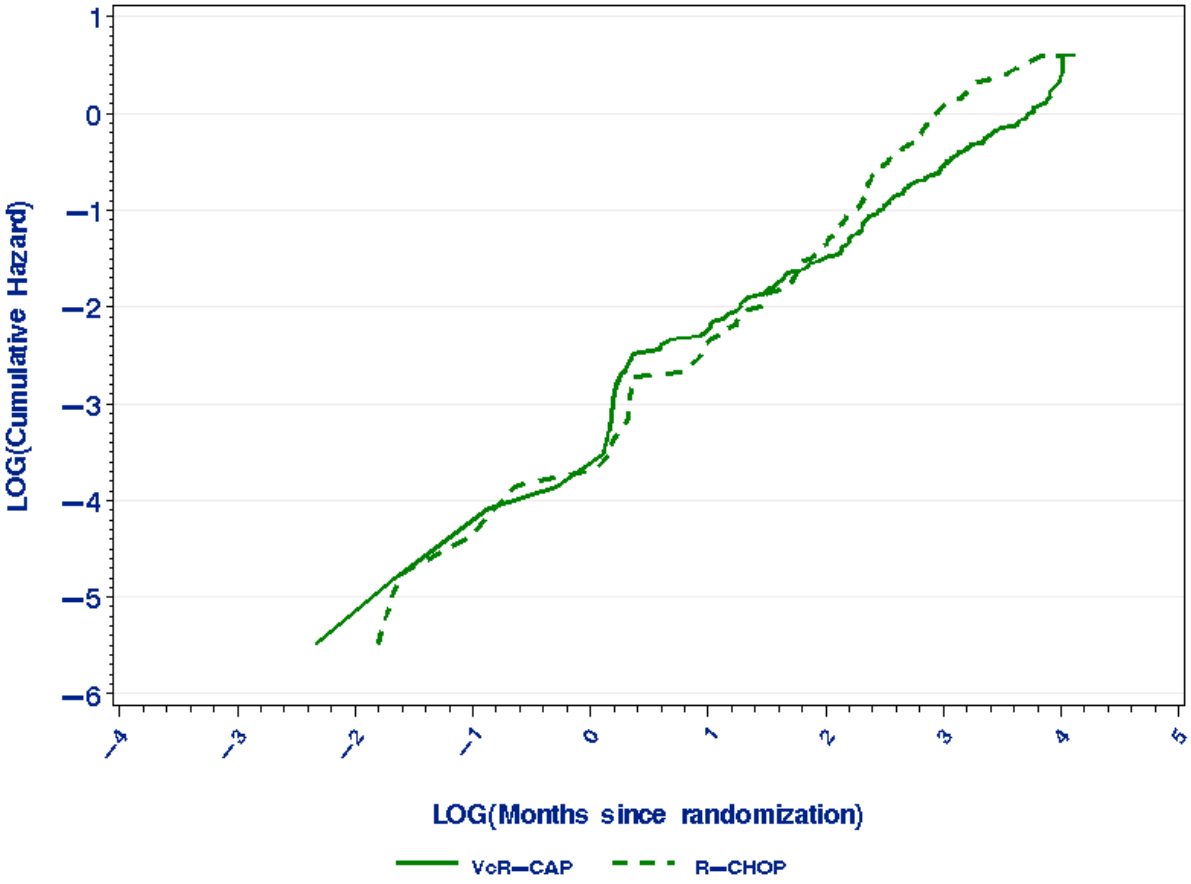
Survival function	Definition	Comments
Progression-free survival (PFS)	Time from start trial to progression or death (whichever occurs first)	In the base-case analysis, an alternative IRC assessment was used to define progression. This alternative assessment is based on the results of sequential assessments instead of only one CT-scan. It is considered more reflective of clinical practice
Overall survival (OS)	Time from start trial to death	Within the model, overall survival is separately estimated for patients with and without progression. For both patient groups, survival is measured from start of the trial to death.
Treatment-free interval (TFI)	Time from last dose of initial treatment to start second-line treatment	Although this is only relevant for patients with progression, it seems as if patients without progression are also include in the estimation as it is measured in all patients who received at least one dose of medication.
Key: IRC = independent review committee; PFS = progression-free survival; OS = overall survival; TFI = treatment-free interval		

Progression-free survival

The definition of progression in the cost-effectiveness analysis differed from the definition used for PFS as primary outcome of the trial. For the latter, progression was evaluated by an independent review committee (IRC) according to the results of a single CT scan at one point in time. However, clinicians indicated that in clinical practice, progression is always confirmed by sequential assessments instead of a single CT scan. Therefore, an alternative assessment of progression according to sequential assessments was used in the economic analysis. These assessments were performed by the IRC. Two scenario analyses were performed for the assessment of progression, including the primary assessment of the IRC and the assessment of progression by the investigator. The primary assessment resulted in an increase of the ICER (£21,766 versus £20,362), while the investigator assessment resulted in a decrease of the ICER (£18,973 versus £20,362).

Individual (i.e. stratified) parametric survival models were estimated for PFS in the two treatment arms, because the log cumulative hazard plots of the two treatment lines were not parallel (Figure 6.2).

Figure 6.2: Log cumulative hazard plots for PFS

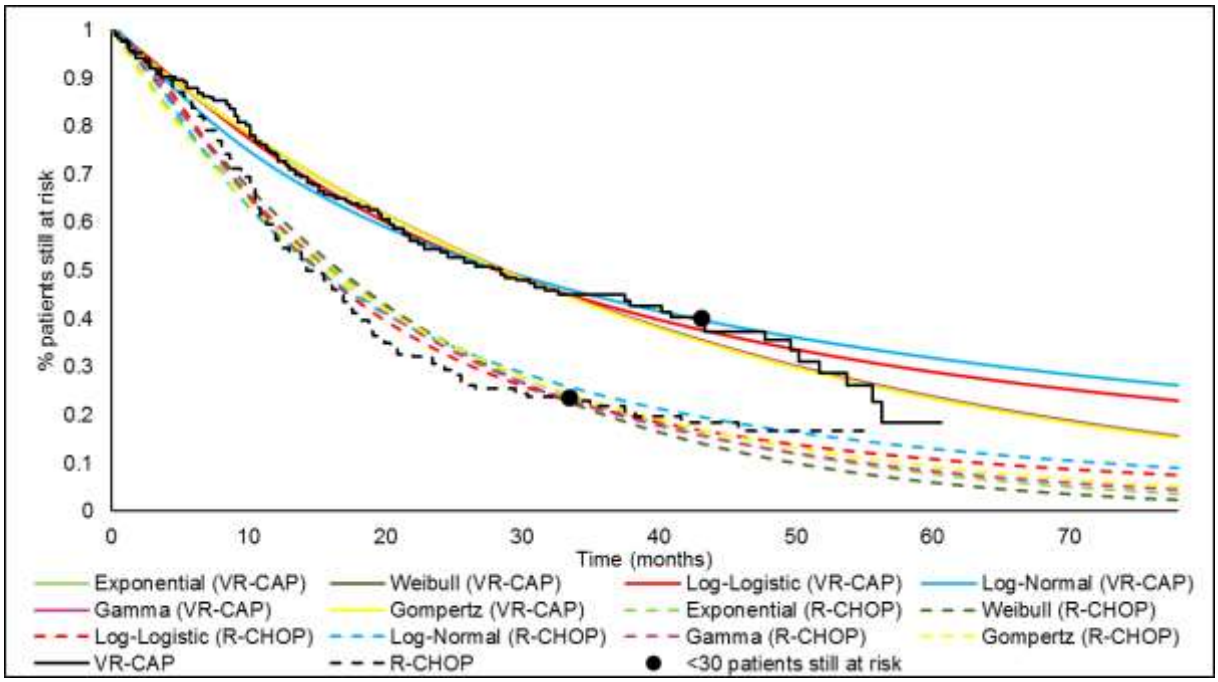


Key: PFS, progression-free survival; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone.

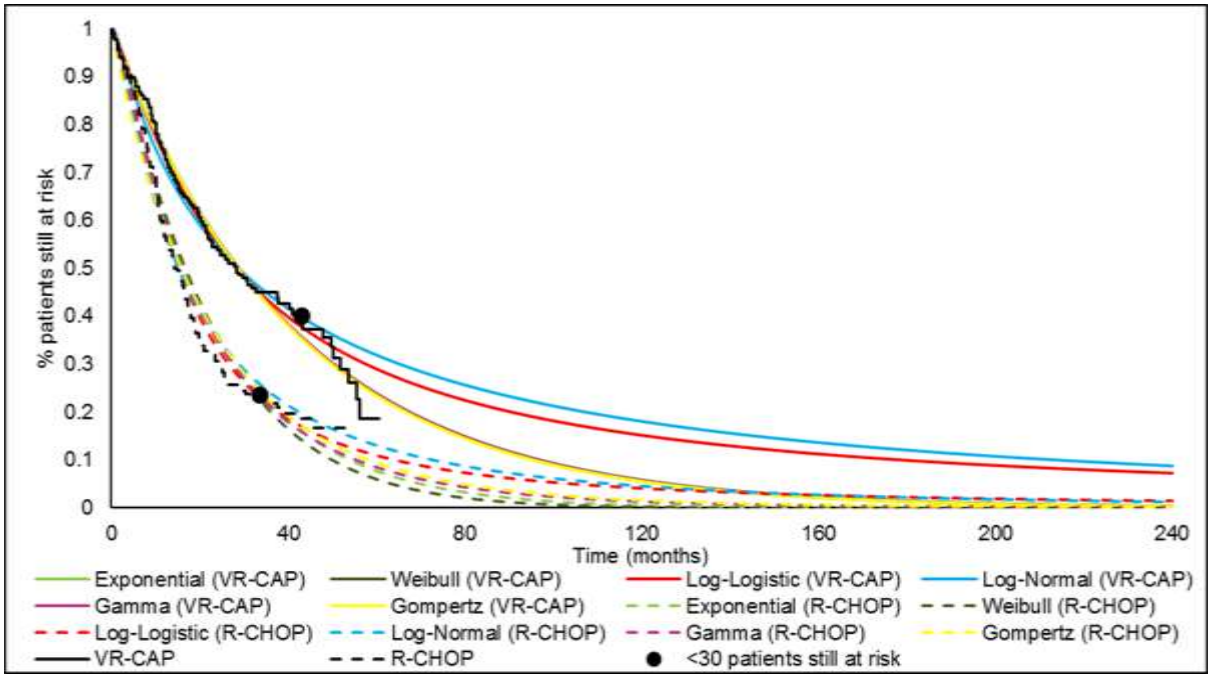
Both the exponential and log-logistic distribution had a (relatively) good fit according to Akaike information criterion (AIC) and Bayesian Information Criterion (BIC) (see Table 35 of the CS[2]) and visual inspection (Figure 6.3). Despite the use of individual survival functions, the same distribution was selected for both treatment arms. Higher PFS in the tail was preferred, because clinical experts had the impression that the observed PFS in the trial was shorter than expected. Therefore, the log-logistic curve was selected for both treatment arms.

Figure 6.3: Fit of parametric models to PFS

A. Short-term duration



B. Long-term duration



Key: PFS, progression-free survival; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone.

Overall survival

It was considered that the extrapolation of OS for all patients resulted in implausible survival estimates due to relatively immature survival data. Subsequently, a literature search was performed to identify a surrogate outcome for OS by identifying relationships between OS and any other clinical relevant outcome, e.g. progressed or not, time to progression, response

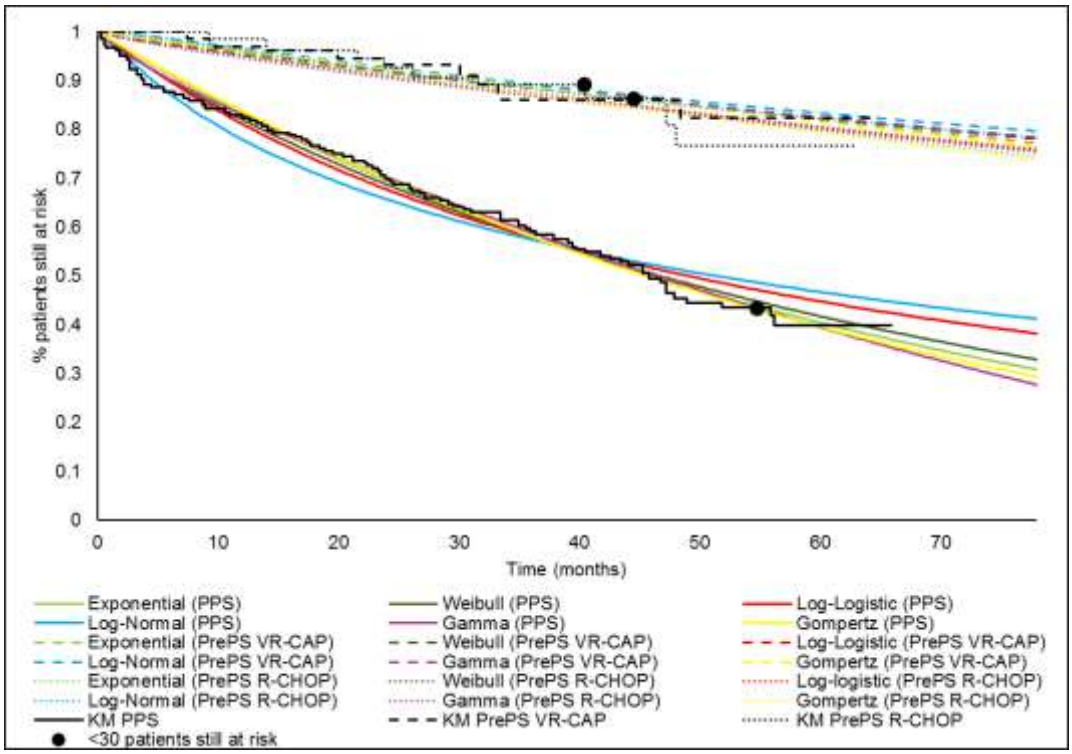
and response duration. A relationship between progression and OS was found in one study[45] and confirmed by UK clinicians. On this basis, it was decided to separately model OS for patients with and without progression. Furthermore, it was assumed that initial treatment only impacts OS for patients without progression, because OS for patients with progression did not differ between the R-CHOP and VR-CAP arm in the LYM-3002 trial. Consequently, OS was estimated for three different groups in the base-case analysis:

- OS for patients without progression (PrePS) treated with R-CHOP
- OS for patients without progression (PrePS) treated with VR-CAP
- OS for patient with progression (PPS) (irrespective of initial treatment)

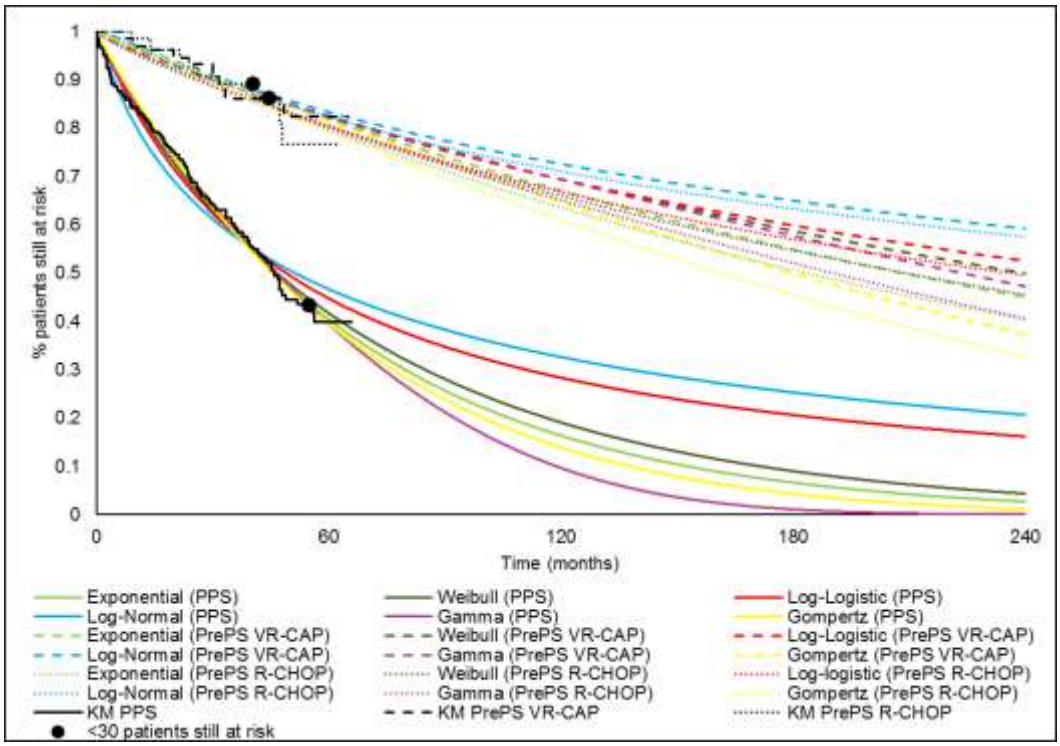
The survival for these three groups are derived from one survival model with 'non-progressed R-CHOP' and 'non-progressed VR-CAP' as covariates in the model. A further distinction of PPS by treatment arm was evaluated in a scenario-analysis by adding "progression VR-CAP" as a covariate in the model. According to statistical (see Table 33 of the CS[2]), visual and long-term fit (Figure 6.4), the exponential distribution was considered as the most reliable distribution to extrapolate OS. Nevertheless, clinicians considered the PrePS as quite high. This overestimation was probably a consequence of the relative immature survival data that did not adequately captured non-disease-specific mortality. Therefore, background mortality was added to the estimated OS for patients without progression. The inclusion of background mortality was conservative, because the ICER decreased by £5,334 if background mortality was excluded. Thereby, it was assumed that all deaths reported in the trial were due to the MCL.

Figure 6.4: Fit of parametric curves to OS

A. Short-term duration



B. Long-term duration



Key: KM, Kaplan–Meier; OS, overall survival; PrePS, pre-progression survival; PPS, post-progression survival; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone.

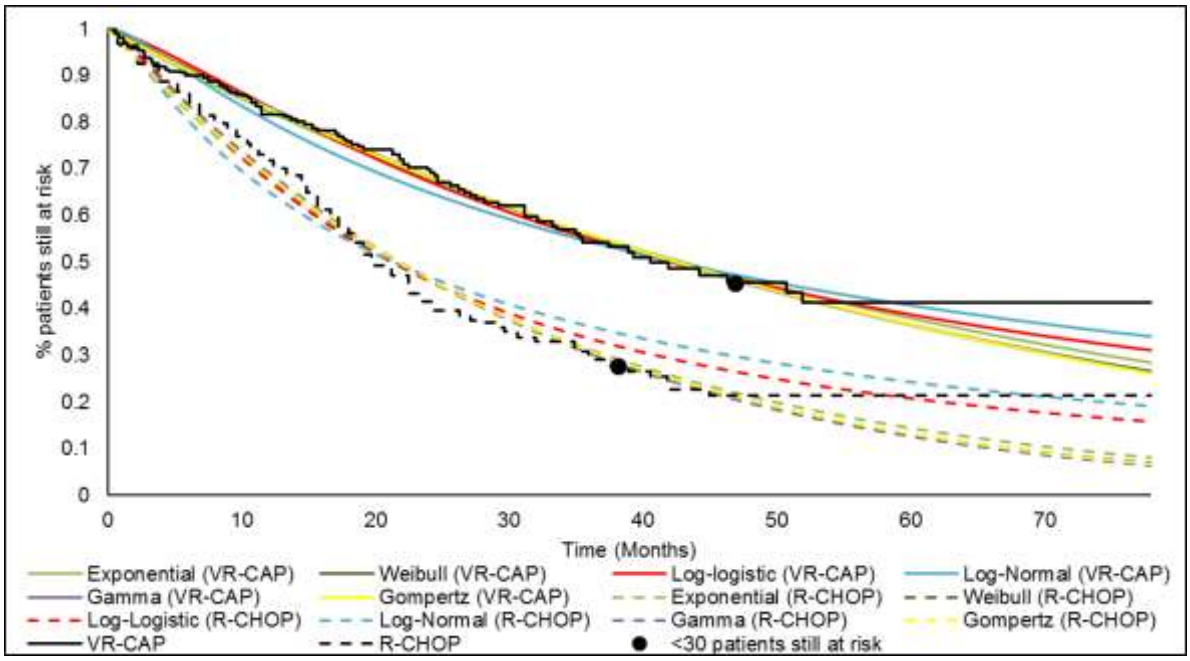
Treatment-free interval

The treatment-free interval (TFI) is defined as the day after the last dose of first line treatment until the start date of the second line treatment. TFI is measured in the safety population, which constitutes of patients who received at least one dose of study medication. Consequently, both patients with and without disease progression are included in the estimation of the TFI. Different parametric survival models were estimated for the TFI with the best fit for the exponential distribution (see Table 75 in Appendix 18 of the CS[15] and Figure 6.5). Initial treatment (VR-CAP versus R-CHOP) was included as a covariate in these survival analyses, thereby allowing different TFIs for the two treatments.

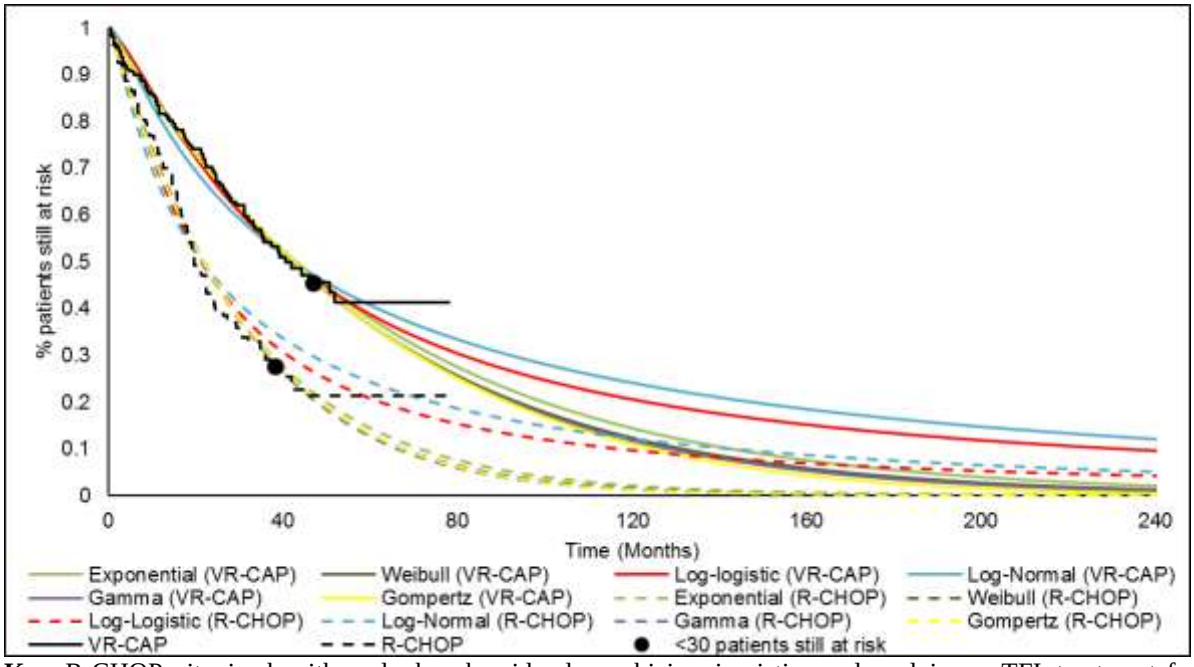
Since TFI starts at the end of the study medication, it is implemented in the model after the average treatment duration. The company stated in the clarification letter that this approach was selected because treatment is administered for a maximum of eight cycles and not continued until progression. Therefore, the TFI already starts in the health state PFS from first-line treatment. The exponential survival function is used to estimate the proportion of patients that start second line treatment. This proportion is corrected for the fact that TFI is calculated in all patients who started with study medication irrespective of disease progression while second line treatment is only administered to patients with disease progression.

Figure 6.5: Fit of parametric models to TFI

A. Short-term duration



B. Long-term duration



Key: R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisone; TFI, treatment-free interval; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisone.

Second-line treatment

Type of second-line treatment differed slightly after R-CHOP and VR-CAP as initial treatment (Table 6.14). As these differences in second-line treatment may impact survival, a regression analysis was performed to assess whether type of second line treatment is significantly associated with survival. All types of second-line treatment were classified in five categories based upon advice of clinical experts. A Cox proportional hazard model was performed, including baseline patient characteristics to correct for confounding. The results

of the regression analyses are presented in a forest plot in figure 6.6. Only the category 'R+chemotherapy' had a significant impact on survival. Non-significant effects of other second-line treatments, like experimental treatments, may be caused by the low number of patients. Despite the significant impact of 'R-Chemotherapy' on survival and the difference in proportion of patients receiving R+chemotherapy after VR-CAP or R-CHOP, survival was not corrected for the type of second-line treatment. The company stated that a conservative approach was used because the proportion of patient receiving R+chemotherapy was slightly in favour of R-CHOP.

The mean duration of second line treatment and PFS from second line treatment were derived from the LYM-3002 trial. It was assumed that these estimates were similar for R-CHOP and VR-CAP (second-line treatment duration = 90 days, PFS 2nd line = 231 days). These average durations are translated to a weekly probability to estimate the probability of going off treatment and progression from 2nd line treatment with the following formula:

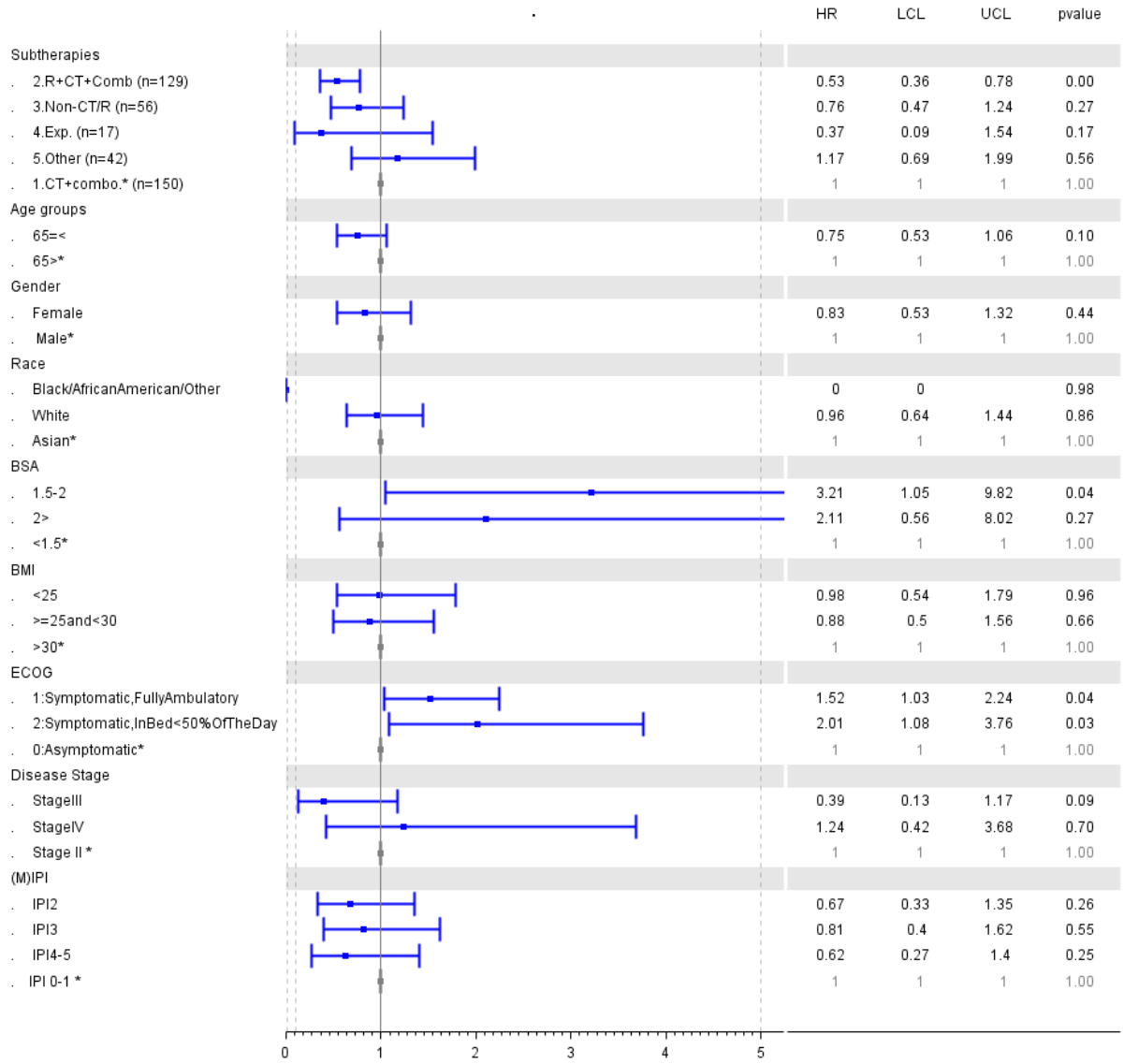
$$\exp\left(-1/\frac{average\ duration\ (in\ days)}{7}\right).$$

Other comparators

As a consequence of the limited evidence from the indirect comparison between VR-CAP and R-bendamustine and R-FC (CS section 4.10[2]), it was assumed by the company that R-bendamustine and R-FC had similar efficacy to R-CHOP. This assumption is conservative for the comparison with R-FC, because OS seems to be worse after R-FC due to limited options for second-line treatment. The assumption of equal efficacy of R-bendamustine was confirmed by clinicians and evidence from a non-randomised study. That study was a large-scale observational study that provided survival data for a number of chemotherapy induction regimens ± rituximab[46]). Only one parameter differed for R-FC and R-bendamustine compared to R-CHOP, which is the average duration on first line treatment. The reason for these differences is the longer duration of a treatment cycle (28 instead of 21 days). Consequently, the average treatment duration is longer and therefore, second line treatment starts at a later point in time. This difference resulted in more QALYs for R-FC and R-bendamustine as compared to R-CHOP. The total life years did not differ between the three treatment arms.

A scenario analysis was performed regarding the inclusion of R-maintenance after first-line treatment. Only two studies were available which could be used to estimate the impact of maintenance treatment on OS and PFS (see Table 40 of the CS[2]). However, as the studies are not directly applicable to the study population or do not provide a direct comparison with no maintenance, different scenarios have been performed to assess the impact of maintenance treatment on the ICER. Within these scenarios, hazard ratios (HRs) are used to adjust the OS and PFS for patients who responded to first-line treatment. Table 6.6 shows the HRs used in the scenarios. These adjustments were only applied during maintenance treatment (two year) which is applied after first-line treatment. Since the hazard ratios are subject to uncertainty, several threshold analyses have been performed with varying HRs (see Section 6.2.11).

Figure 6.6: Regression analysis OS from start of subsequent treatment



Key: Chemo, chemotherapy; Combo, combination therapy; Exp, experimental, HR, hazard ratio, n, number; OS, overall survival; R, rituximab.

Table 6.6: Scenario analyses regarding the effect of maintenance treatment

	Scenario 1	Scenario 2
Source	NICE TA226[47]	European MCL Elderly trial[25, 47]
Hazard ratio for OS	1	0.48
Hazard ratio for PFS	0.55	0.41
Key: MCL = Mantle Cell Lymphoma; OS = overall survival; PFS = progression-free survival		

ERG comment: In general, the ERG has several comments on the way the PFS, OS and TFI were estimated. The company followed the approach outlined in the DSU report on the extrapolation of survival curves[48] and in general reasonable choices were made. The ERG does however have some comments.

The ERG has some concerns regarding the estimation of treatment effectiveness in the model. First, it is questionable whether the best fitted model was selected to model the PFS for the VR-CAP arm. In the base-case analysis, the log-logistic distribution was selected for both treatment arms based upon clinical expert opinion. However, the exponential distribution showed the best statistical fit for the VR-CAP arm. The company did report the results of a scenario-analysis in which the exponential distribution was used to model both treatment arms. That scenario resulted in an increase of the ICER to £28,113. However, it can be argued that this approach is still suboptimal as the exponential distribution is not the most optimal distribution to model PFS for the R-CHOP arm. Since individual models were estimated for both treatment arms, it is feasible to use different distributions to model the PFS in the two treatment arms. Therefore, the ERG performed an additional analysis in which the exponential distribution was used to model the PFS in the VR-CAP arm and the log-logistic distribution to model the PFS in the R-CHOP arm. This was included in the ERG base case as presented in Section 6.3. Another concern regarding the estimation of PFS is the use of a stratified model (i.e. individual models for the two treatment arms) as the evidence for non-parallel log-logistic curves is debatable due to the crossover of the curves at the end of follow-up. Therefore, the ERG performed an exploratory analysis including the unstratified model for PFS (see Section 6.3).

The ERG agrees that immature data may bias the extrapolation of survival data. However, no reasonable motivation has been provided by the company for the implausible results of the extrapolation. In the company's response to the clarification letter, the company compared the extrapolated survival curves with the observed survival in 1,151 Swedish and Danish patients who have been treated with rituximab based chemotherapy.[46] Although the included patients and treatments in this study were not directly comparable with the LYM-3002 trial, it seems that a Weibull distribution had the best fit with the data from Abrahamsson et al[46] Therefore, the company performed the cost-effectiveness analyses using OS based upon the Weibull distribution instead of PrePS and PPS. This approach resulted in an increase of the life years after R-CHOP (5.81 versus 5.69) and a decrease of the life years after VR-CAP (6.42 versus 6.44). Consequently, the ICER per QALY gained increased from £20,362 to £22,052. The company stated that the approach with the distinction in survival is preferred, as there is no clinical evidence that survival for patients with progression differs between VR-CAP and R-CHOP.

However, if data are too immature to model overall survival for all patients, it is questionable whether sufficient data is available to separately estimate long-term survival for patients with and without progression. This distinction would reduce the total number of patients at risk, and thereby may increase the uncertainty about the long-term survival. In the response to the clarification letter, the company stated that uncertainty is reduced for patients with progression as a smaller proportion of patients at risk is still alive at time of evaluation. Furthermore, they claim that the data of the two treatment arms is pooled and thereby the total sample size is increased. However, the company did not reflect on the uncertainty of survival for patients without progression (PrePS). For this patient group, only a minority of patients died during the study period and data has not been pooled. Nevertheless, the

company justifies the use of different survival for patients with and without progression by referencing one study in MCL and one study in NHL in which better PFS is associated with better OS[45, 49] Another concern regarding the modelling of survival is the assumption that survival for patients without progression differs between treatment arms. It can be argued that due to the immature data, it is not feasible to identify any differences in OS between treatment arms. This is also supported by observed KM-curves for the treatment arms (Figure 6.4A). A more conservative approach is to assume no treatment effect on OS, but only on PFS. The ERG has performed this additional analysis (see Section 6.3).

The TFI is measured in all patients who received at least one dose of study medication, irrespective of treatment response. However, as second-line treatment is only administered to patients with disease progression, all patients without progression are likely to be censored in this analysis. Consequently, censoring is non-random and it is expected that the treatment-free interval is overestimated. This overestimation may be larger for the VR-CAP arm since PFS is better and therefore more patients are censored. In order to avoid this bias in the estimation of the TFI, the company was asked in the clarification letter to restrict the analysis of the treatment-free interval to patients with progression. The company did not provide the estimate of the TFI restricted to patients with progression, but stated that it was corrected for in the model by dividing the proportion of patients that would start second-line treatment based on the TFI curves by the proportion of patients that have progressed. However, this does not correct the TFI appropriately. Additionally, the ERG found that the proportion of patients that have transited from progressed first line to second line treatment was incorrect in the model as it was estimated from all previous cycles since the start of the model instead of the preceding cycle only (leading to an underestimation of the TFI). The correction of this error (by selecting the preceding cycle only) was not feasible, because at some point in time the proportion of patients with progression was very small, resulting in probabilities to start second line treatment that were larger than one.

Another minor concern regarding the TFI is that it does not take into account that patients may progress before the end of the full treatment schedule (eight cycles). For these patients, it is expected that the TFI starts at time of progression instead after the average first-line treatment duration. Furthermore, the average treatment duration is slightly overestimated because it is estimated from patients who did not stop treatment due to progression or death. Consequently, the start of the TFI is postponed for a subset of the patients. However, since only a small proportion of patients progress before the end of the full treatment schedule and since this proportion is smaller for patients treated with VR-CAP, it is expected that the company's approach has only marginal impact on the ICER and is conservative.

Due to the various concerns regarding the estimation of the TFI, the ERG decided that it would be more reliable to neglect a TFI and implement second-line treatment at time of progression. This approach is more conservative, because the observed TFI in the LYM-3002 is twice as long after VR-CAP compared to R-CHOP. Nevertheless, the ERG is aware that a TFI can be common in lymphoma-type diseases. Therefore, two scenario analyses have been performed based upon assumptions regarding the time from progression to start second-line treatment. In one scenario analysis, it is assumed that the time to second-line treatment does

not differ between R-CHOP and VR-CAP. T The other scenario analysis assumes that the time to second-line treatment is twice as long after VR-CAP compared to R-CHOP (12 versus six months).

The ERG explored the impact of the assumption of equal PPS and did agree with the company that a conservative approach was used by not correcting survival after second line treatment for type of second-line treatment.

6.2.7 Adverse events

The economic model includes all adverse events (AEs) with a grade ≥3 and incidence above 5.0% in the VR-CAP or R-CHOP arm as observed during first-line treatment in the LYM-3002 trial (i.e. grade ≥3 AE that were reported in at least 5% of the patients in either study group during the treatment period).[9] Grade 2 peripheral sensory neuropathy and grade 3 or higher alopecia and sepsis were also included in the model because these were deemed of clinical significance by the medical specialists attending the UK advisory board.[7] AE were incorporated separately for grade 3 and grade 4/5. The AE incidences for VR-CAP and R-CHOP were retrieved from the LYM-3002 trial[9] while the AE incidence for R-FC was retrieved from a study by Kluin-Nelemans et al (as-treated analysis) including patients 60 years of age or older with mantle-cell lymphoma stage II to IV[25] and the AE incidence for R-bendamustine was retrieved from a study by Flinn et al[24] including treatment-naive patients with indolent non-Hodgkin's lymphoma or mantle cell lymphoma.

The weekly AE probability was obtained by firstly calculating the annual AE rate for each included AE separately for both treatment arms (based on the number of events in the LYM-3002 trial and the total patient years on treatment). Subsequently, these annual rates were converted to weekly probabilities for each AE.

As the AE incidences retrieved for R-FC and R-bendamustine were not reported separately for grade 3 and grade 4/5, the same distribution for grade 3 and grade 4/5 was assumed as for R-CHOP. Moreover, in case specific AE incidence was not available for R-FC and R-bendamustine, an incidence of zero was assumed.

Duration of AE's was incorporated as treatment specific for VR-CAP and R-CHOP (retrieved from the LYM-3002 trial[9]) while for R-FC and R-bendamustine this was assumed to be equal as for R-CHOP. See Table 6.7 for an overview of the AEs probabilities and AE durations incorporated in the model.

No AE for second-line treatment were incorporated in the model.

Table 6.7: AEs weekly probabilities and duration
(Based on Table 45 of the CS[16])

Adverse event	VR-CAP		R-CHOP		R-FC		R-bendamustine		Duration (weeks)	
	Grade 3	Grade 4/5	Grade 3	Grade 4/5	Grade 3	Grade 4/5	Grade 3	Grade 4/5	VR-CAP	R-CHOP[c]
Neutropenia	1%	4%	1%	3%	0%[b]	0%[b]	0%[b]	0%[b]	1.09	1.30
Thrombocytopenia	1%	2%	0%	0%	0%	0%	0%[b]	0%[b]	1.04	1.44
Anaemia	1%	0%	1%	0%	0%	0%	0%[b]	0%[b]	0.94	1.39
Leukopenia	1%	2%	1%	1%	1%	1%	0%[b]	0%[b]	1.10	1.35
Lymphopenia	1%	1%	0%	0%	1%	0%	0%[b]	0%[b]	1.19	2.39
Febrile neutropenia	1%	0%	0%	0%	0%	0%	0%	0%	0.95	1.19
Pneumonia	0%	0%	0%	0%	0%[b]	0%[b]	0%	0%	2.70	2.29
Fatigue	0%	0%	0%	0%	0%	0%	0%	0%	5.06	3.29
Diarrhoea	0%	0%	0%	0%	0%[b]	0%[b]	0%	0%	0.71	0.89
Peripheral sensory Neuropathy	1%[a]	0%	0%[a]	0%	0%	0%	0%	0%	8.19	21.24
Alopecia	0%	0%	0%	0%	0%[b]	0%[b]	0%	0%	46.43	10.68
Sepsis	0%	0%	0%	0%	0%[b]	0%[b]	0%[b]	0%[b]	1.81	0.50

Key: AE, adverse event; R-bendamustine, rituximab with bendamustine; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; R-FC, rituximab with fludarabine, and cyclophosphamide; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone.
[a] Includes Grade 2 peripheral sensory neuropathy.
[b] Assumed to be zero since data was not available in the literature.
[c] Duration of AEs for R-FC and R-bendamustine was assumed equal to AE durations for R-CHOP.

ERG comment: The AE incidences and durations for VR-CAP and R-CHOP were retrieved from the LYM-3002 trial[9] while AE incidences for R-FC and R-bendamustine were retrieved from the literature[24, 25] with (slightly) different populations. AE durations for R-FC and R-bendamustine were assumed to be equal to those for R-CHOP. The inclusion of AE incidences for R-FC directly based on the literature instead of based on an indirect treatment comparison (Bucher method) seems to be a conservative approach given that the inclusion of indirect treatment comparison decreases the ICERs. However, since the scenario presented by the company also includes indirect treatment comparisons for PFS and OS, the isolated effect using indirect treatment comparisons for AE is unclear (see response to clarification question C15[16]). Moreover, the assumption of treatment dependent AE durations is likely to be conservative (see scenario analysis EFF18 in the CS[2]).

The company stated that the economic model includes all AEs with a grade ≥3 and incidence above 5.0% in the VR-CAP or R-CHOP arm as observed during first-line treatment in the LYM-3002 trial. However, the clinical study report (Appendix TAE09B)[21] showed that the grade ≥3 AE incidences for alopecia and sepsis are not above this threshold and hence the inclusion of these AE by the company is not consistent with their decision rule.

The assumption of zero AE for R-FC and R-bendamustine in case specific AE incidence was not available can be considered conservative.

The model was restricted to AE during first-line treatment. The exclusion of AE during second-line treatment is considered to be conservative as patients spent least time in the second-line when treated with VR-CAP (see Table 6.20) and hence a greater negative impact of AE associated with second-line treatment would be expected for R-CHOP, R-FC and R-bendamustine compared with VR-CAP. This is confirmed by the company (see response to clarification question C14[16]).

Finally, the ERG identified a model error (regarding the calculation of AE cycle probabilities) which is corrected in the ERG base case (see response to clarification question C16 for more details[16]).

6.2.8 Health related quality of life

Base case

Quality of life utilities for the health states in the model (progression-free survival (PFS) from first-line treatment, progressed from first-line treatment, PFS from second-line treatment and progressed from second-line treatment) were derived from different sources, i.e. the LYM-3002 trial, a literature review and clinician input.

The LYM-3002 trial assessed health-related quality of life (HRQoL) in patients with newly diagnosed MCL with advanced disease who are unsuitable for haematopoietic stem cell transplantation (HSCT). EQ-5D measurements were taken at baseline, at day one of every treatment cycle and at the end-of-treatment visit. Health states were valued using the UK-tariff. A mixed model approach was used to estimate mean utility values for the health states used in the model; PFS (0.764; N=3,033) and progressed from first-line treatment (0.693; N=162). According to the company, this method likely underestimates the impact of progression on QoL since utilities were only assessed at short term, with the latest assessment

being at the end-of-treatment visit, which was performed 30 days after the last dose of investigational product was administered.. QoL estimates for the progressed health states are therefore likely overestimated.

As utility data for second-line treatment were not captured by the LYM-3002 trial, a systematic literature review was designed to identify relevant QoL data for patients with MCL. No studies were identified in which utilities of the specific MCL population were reported. Therefore, NHL was considered the best proxy for utility values of MCL patients. Four potentially relevant primary utility studies[41, 50-52] and one mapping study[53] were identified reporting utilities in previously untreated patients with NHL.

Of the five potentially relevant studies, the paper by Doorduijn et al (2005)[41] was considered to be the most suitable for utilisation within the model for MCL since primary EQ-5D utility values were reported by response and progression based health states and the study was conducted in patients with aggressive NHL which was considered to be the most clinically relevant patient population. Utilities for the progressed from second-line treatment (0.45) health state were derived by averaging over low and high age-adjusted Prognostic Indices (aaPI) using the following formula: [aaPI 0-1 Baseline (0.74) + progression (-0.24)] + [aaPI 2-3 Baseline (0.44) + progression (-0.04)] / 2 = 0.45.

In addition to the primary utility and mapping studies, a few of other studies were identified through the literature review i.e. Soini et al 2012[54] , Hayslip et al 2008[55] and Deconick et al 2010[56]. These studies did not report primary utility values but applied utilities in their cost-effectiveness analyses that were considered applicable to the model. However, all these studies were conducted in patients with follicular lymphoma (FL), which is not deemed reflective of the HRQL impact of MCL. These utility values were therefore not used in the base case analysis but explored in scenario analyses in the CS.

The utility value for PFS from second line treatment was assumed to be equal to PFS from first-line treatment (0.764). This assumption was based upon clinician input and previous modelling.

Adverse events

Utilities values for adverse events (AE) of first-line treatment were included in the model and derived from the LYM-3002 trial. Utility decrements were estimated by subtracting the utility value before developing an AE from the utility value with the AE. The utility decrements were applied for the duration of the AE. Utilities for AEs of second-line treatment were not included as no EQ-5D data were available from the LYM-3002 trial.

Summary

A summary of the quality of life values used in the base case and scenario analysis is shown in Table 6.8.

Table 6.8: Summary of QoL values used for CEA

	Health state	Utility value, mean*(95% CI)*	Reference in submission	Source
	Base case			
	Progression-free survival from first-line treatment	0.764*(0.746;0.781)*	Table 42 Section 5.4 p. 127	LYM-3002[21]
	Progressed from first-line treatment	0.693 *(0.639:0.744)*		
	Progression-free survival from second-line treatment	0.764 *(0.746;0.781)*	p. 126	Assumption, clinician input
	Progressed from second-line treatment	0.45 *(0.206;0.708)*	Table 44 p. 166	Doorduijn et al. 2005[41], clinician input (p. 126)
Nr.	**Scenario analysis**			
1.	Progression-free survival from first-line treatment	0.764		LYM-3002[21]
	Progressed from first-line treatment	0.693		LYM-3002[21]
	Progression-free survival from second-line treatment	0.693		Assumption
	Progressed from second-line treatment	0.45		Doorduijn et al. 2005[41]
2.	Progression-free survival from first-line treatment	0.764		
	Progressed from first-line treatment	0.693		
	Progression-free survival from second-line treatment	0.764		
	Progressed from second-line treatment	0.693	Table 63 (no further clarification)	Assumption
3.	Progression-free survival from first-line treatment	0.610	Table 64 & section 5.4 p. 130	Doorduijn et al. 2005[41]
	Progressed from first-line treatment	0.450		
	Progression-free survival from second-line treatment	0.610		
	Progressed from second-line treatment	0.45		
4.	Progression-free survival from first-line treatment	0.764	Table 42 Section 5.4 p. 127	LYM-3002[21]
	Progressed from first-line treatment	0.604	Table 64 & table 44	Soini et al. (2012): 0.764-(0.78-0.62)[54]
	Progression-free	0.764	p. 126	Assumption

	Health state	Utility value, mean*(95% CI)*	Reference in submission	Source
	survival from second-line treatment			
	Progressed from second-line treatment	0.45	Table 64 & table 44	Doorduijn et al. 2005[41]
5.	Progression-free survival from first-line treatment	0.764	Table 42 Section 5.4 p. 127	LYM-3002[21]
	Progressed from first-line treatment	0.694	Table 64 & table 44	Hayslip et al. (2008): 0.764-(0.805-0.618)[55]
	Progression-free survival from second-line treatment	0.764	p. 126	Assumption
	Progressed from second-line treatment	0.45	Table 64 & table 44	Doorduijn et al. 2005[41]
6.	Progression-free survival from first-line treatment	0.764	Table 42 Section 5.4 p. 127	LYM-3002[21]
	Progressed from first-line treatment	0.577	Table 64 & table 44	Deconinck et al. 2010[56]
	Progression-free survival from second-line treatment	0.764	p. 126	Assumption
	Progressed from second-line treatment	0.45	Table 64 & table 44	Doorduijn et al. 2005[41]
	Exploratory comparison			
	Progression-free survival from first-line treatment	0.764		
	Progressed from first-line treatment	0.693		
	Progression-free survival from second-line treatment	0.764		
	Progressed from second-line treatment	0.45		

Table 6.9: Disutilities for adverse events

Adverse events			
Adverse event of first-line treatment	**Utility decrement**	**Reference**	
Neutropenia	-0.000	Table 46	LYM-3002[21]
Thrombocytopenia	0.010		
Anaemia	-0.009		
Leukopenia	-0.023		
Lymphopenia	-0.026		
Febrile neutropenia	0.090		
Pneumonia	-0.059		
Fatigue	-0.087		
Diarrhoea	-0.000		
Peripheral sensory neuropathy	-0.098		
Sepsis	-0.171		

Adverse events			
Adverse event of first-line treatment	**Utility decrement**	**Reference**	
Alopecia	-0.130		

ERG comment: The ERG identifies substantial uncertainty regarding the utilities for the health states: progression from first-line treatment, PFS from second-line treatment and progression from second-line treatment. The utility for progression from first-line treatment is derived from the LYM-3002. However, the company already observed that it might be an overestimation, because all patients were still on treatment while reporting their quality of life. In their response to the clarification letter, the company acknowledged the uncertainty regarding this estimate and performed some scenario analyses with different utility values for this health states. These scenario analyses had only minor impact on the ICER, because the total duration in this health state was short. In the ERG base case, the treatment-free interval is not included. Therefore, patients move directly from the PFS from first-line treatment to PFS from second-line treatment. Consequently, the utility for progression from first-line treatment is no longer relevant.

The company assumed similar utility for PFS from first and second-line as clinicians report an improved quality of life if second-line treatment is initiated after a progression of the disease. Although quality of life can be improved once treatment is reinitiated, it is not by definition equal to the quality of life in patients with PFS from first line treatment. The ERG considers that the quality of life utility is overestimated for this health states. Nevertheless, this assumption is conservative because fewer patients in the VR-CAP arm receive second-line treatment in comparison with the other treatments.

As mentioned above, the utility value used in the base case model for progression from second-line treatment is 0.45. This value is estimated by applying a disutility for progressed disease to the baseline utility as observed by Doorduijn et al (2014).[41] However, the evidence from the Doorduijn study is rather weak as disutilities are only based on 26 observations. Therefore, the company was asked in the clarification letter (Section C – question 8c) to provide further justification of using a utility of 0.45 that is significantly lower compared to the value used for progressed from first-line treatment (0.693). The company justifies using the 0.45 for progressed from second-line treatment with clinical expert opinion who stated that patients who have progressed from second-line treatment have a worse quality of life than patients who have progressed from first-line treatment because they have limited treatment options still available and, moreover, their disease and symptoms will have worsened over time. The estimate of 0.45 by Doorduijn et al,[41] was seen as a sensible estimate. The company acknowledge the fact that this utility was measured in NHL patients and that only a few observations were available at time of progression. However, the company stated that this is the best available evidence given the advice from clinical experts.

The ERG does not agree with using a utility of 0.45 in the base-case analysis given the uncertainty of this estimate. Furthermore, the utility is calculated by subtracting a disutility from a baseline utility in NHL patients. Since the baseline utility for MCL patients has been measured in the LYM-3002 trial (utility for patients with PFS from first line), the ERG

proposes to subtract the disutility as observed by Doorduijn et al (2014) from this baseline utility (0.764). This methodology has also been applied by the company for the scenario analyses with data from other studies. Two different disutilities were reported in Doorduijn et al (2014): one for patients with an age-adjusted International Prognostic Index (aaPI) of 0-1 and one for patients with an aaPI of 2-3 (disutility of 0.24 and 0.04, respectively). Although the aaPI was not available for the LYM-3002 trial, another risk classification showed that about 50% of the patients have a low risk. Therefore, in the ERG base-case analysis, the average disutility of the two risk group (0.14) is subtracted from the baseline utility in the LYM-3002 trial (0.765) to estimate quality of life utility for progression from second-line treatment. The resulting utility for this health state is therefore 0.624. An additional scenario analysis with the disutility for the low risk group (aaPI 0-1) was applied because the low disutility for the high risk group might be a consequence of the low baseline utility in this group.

The ERG noted that some utility values used for the secondary scenario analyses are not correctly entered into the model of the company. As described in the clarification letter (Section C – question C9a), the correct utility values for the scenario analysis based on Soini et al (2012)[54] and Hayslip et al (2008)[55] for progressed form first-line treatment should be 0.604 and 0.694, respectively. However, utilities entered in the model are 0.924 and 0.834. The ERG corrected these errors and ran the model again. The ICERs for VR-CAP versus R-CHOP for these two scenario analysis increased from £20,083 to £20,472 (scenario analysis based on Soini et al) and £20,191 to £20,361 (scenario analysis based on Hayslip et al).

Finally, the ERG noted that for two adverse events the disutility was positive (thrombocytopenia and febrile neutropenia), i.e. having an adverse event led to a utility gain in those patients. In the clarification letter, the ERG asked the company why these disutilities could be positive. The company responded (see response to clarification question C11b[16]) that the utility decrements were most likely positive for some adverse events due to the low number of measurements during these adverse events. In light of this, the company suggested that these values are not robust and should not have been applied to the model. A scenario analysis where no utility decrements at all are applied for AEs was included in the company submission (QOL 7 in Table 72 of the CS) and resulted in an ICER of £20,340 indicating that applying utility decrements for adverse events had very little impact on the ICER.

6.2.9 Resources and costs

Categories considered for resource use and costs were: drug acquisition and administration costs, health state costs, adverse event costs and end of life costs.

Drug acquisition and administration costs

Drug wastage was accounted for in the calculation of drug acquisition and administration costs (i.e. it was assumed that patients received only whole vials and there was no vial sharing). The unit prices for first- and second-line treatments can be found in Table 48 of the CS (retrieved from the electronic market information tool (eMit) and monthly index of medical specialities (MIMS)).[2] The average number of vials/units required per administration and cost per administration for separate drugs are calculated based on an average body surface area of 1.91m^2 (average baseline BSA for patients from the US, Canada and Western

European in the LYM-3002 trial[21]) and presented in Table 6.10. The number of administrations per treatment cycle, maximum number of treatment cycles and costs per treatment cycle for the first and second-line treatments are presented in Table 6.11. The duration for second-line treatment was set to 90 days (see Table 39 of the CS[2]).

Administration costs were £245 per intravenous (IV) administration and £136 per oral administration. For first-line treatment that included an orally administered drug, one additional oral administration visit was assumed at the start of treatment. For second-line, no oral administration visits were assumed for simplicity.

Table 6.10: Number of vials per administration, number of administrations per week and costs per administration
(Partly based on table 50 of the CS[2])

Intravenous therapies			
Drug	**Vial size**	**Vials per administration[a]**	**Cost per administration**
Bortezomib	3.5 mg	1	£762.38
Rituximab	100 mg	2.66	£1,338.09
	500 mg	1	
	1400 mg	0	
Cyclophosphamide	500 mg	1.33	£28.27
	1000 mg	1	
Doxorubicin	10 mg	0.05	£8.91
	50 mg	2	
Vincristine	1 mg	2	£6.17
	2 mg	0	
Bendamustine	25 mg	3.38	£510.28
	100 mg	1	
Cytarabine	100 mg	3.7	£40.37
	500 mg	1	
	1000 mg	1	
	2000 mg	1	
Temsirolimus	30 mg	3	£1,860.00
Oral therapies			
Drug	**Unit size**	**Units per administration[a]**	**Cost per administration**
Prednisolone	5 mg	20	£0.26
Fludarabine	10 mg	7	£141.16
Cyclophosphamide	50 mg	1	£0.82
Ibrutinib	140 mg	4	£204.40

Intravenous therapies			
Drug	**Vial size**	**Vials per administration[a]**	**Cost per administration**
Lenalidomide	10 mg	1	£180.00

[a] The vials per administration and units per administration are averages for the whole population (including wastage) and independent on the comparator (e.g. Doxorubicin dosing is equal for VR-CAP and R-CHOP).

Table 6.11: Dosing regimens, drug acquisition costs and administration costs
(Based on tables 51 and 52 of the CS[2])

Regimen	Drug	Dose	Treatment cycle length (days)	Max number of treatment cycles	Admins per treatment cycle	Admin type	Drug costs per treatment cycle	Administration costs per treatment cycle
Treatment provided during the first-line only								
VR-CAP	Bortezomib	1.3mg/m^2	21	6 (or 8 if first response in cycle 6)	4	IV	£4,426	Cycle 1: £1,116 Other cycles: £980
	Rituximab	375mg/m^2			1	IV		
	Cyclophosphamide	750mg/m^2			1	IV		
	Doxorubicin	50mg/m^2			1	IV		
	Prednisone	100mg			5	Oral		
Treatments provided during and after the first-line								
R-CHOP	Rituximab	375mg/m^2	21	6 (or 8 if first response in cycle 6)	1	IV	£1,383	Cycle 1: £381 Other cycles: £245
	Cyclophosphamide	750mg/m^2			1	IV		
	Doxorubicin	50mg/m^2			1	IV		
	Vincristine	1.4mg/m^2			1	IV		
	Prednisone	100mg			5	Oral		
R-FC	Rituximab	375mg/m^2	28	8	1	IV	£1,764	Cycle 1: £381 Other cycles: £245
	Fludarabine	70mg			3	Oral		
	Cyclophosphamide	50mg			3	Oral		
R-bendamustine	Rituximab	375mg/m^2	28	6	1	IV	£2,359	£490
	Bendamustine	90mg/m^2			2	IV		

Regimen	Drug	Dose	Treatment cycle length (days)	Max number of treatment cycles	Admins per treatment cycle	Admin type	Drug costs per treatment cycle	Administration costs per treatment cycle
Treatments provided after the first-line only								
Ibrutinib	Ibrutinib	560 mg	7	N/A	7	Oral	£1,431	£0
R-cytarabine	Rituximab	375 mg/m^2	21	N/A	1	IV	£1,500	£490
	Cytarabine	200 mg/m^2			4	IV		
Bendamustine	Bendamustine	90 mg/m^2	28	N/A	2	IV	£1,021	£490
Cytarabine	Cytarabine	2000 mg/m^2	21	N/A	4	IV	£161	£490
Rituximab	Rituximab	375 mg/m^2	21	N/A	1	IV	£1,338	£245
Lenalidomide	Lenalidomide	10 mg	28	N/A	21	Oral	£5,040	£0
Temsirolimus	Temsirolimus	75 mg	7	N/A	1	IV	£1,860	£245
CHOP	Cyclophosphamide	750 mg/m^2	21	N/A	1	IV	£45	£245
	Doxorubicin	50 mg/m^2			1	IV		
	Vincristine	1.4 mg/m^2			1	IV		
	Prednisone	100 mg			5	Oral		
Vincristine	Vincristine	1.4 mg/m^2	21	N/A	1	IV	£6	£245

Key: CHOP, cyclophosphamide, doxorubicin, vincristine and prednisolone; IV, intravenous; N/A, not applicable; R-bendamustine, rituximab with bendamustine; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; R-cytarabine, rituximab with cytarabine; R-FC, rituximab with fludarabine and cyclophosphamide; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone

The average first-line treatment costs in the economic model are reduced since 1) not all patients receive treatment during all treatment cycles and; 2) not all patients received the full per-protocol dose (e.g. due to dose modifications in response to AEs).

The percentages of patients receiving first-line treatment cycle are presented per treatment in Table 6.12. For VR-CAP and R-CHOP this was based on the LYM-3002 trial.[9] For R-FC and R-bendamustine these percentages were assumed to be equal as R-CHOP.

The drug costs per treatment cycle presented in Table 6.11 were based on per protocol dosing regimens. These drug costs were adjusted using the relative dose reductions presented in Table 6.12 to reflect the average dose received in the LYM-3002 trial.[9] For R-FC and R-bendamustine no dose reductions were assumed.

Table 6.12: Percentage of patients receiving first-line treatment per treatment cycle
(Based on Table 53 of the CS[2])

	% receiving cycle	
Cycle	**VR-CAP**	**R-CHOP**
1	100%	100%
2	98%	97%
3	95%	94%
4	90%	90%
5	88%	86%
6	85%	84%
7	14%	17%
8	13%	17%

Key: CSR, clinical study report; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone

.

Table 6.13: Difference between average dose received and per protocol dose per cycle in the LYM-3002 trial
(Based on Table 54 of the CS[2])

VR-CAP					
Cycle	**Bortezomib**	**Rituximab[a]**	**Cyclophosphamide**	**Doxorubicin**	**Prednisolone**
1	11%	0%	0%	0%	4%
2	13%	0%	5%	2%	4%
3	14%	0%	7%	3%	4%
4	18%	0%	9%	4%	5%
5	22%	-0%	10%	6%	6%
6	24%	0%	11%	7%	6%
7	26%	0%	9%	4%	2%
8	26%	0%	12%	3%	2%
R-CHOP					
Cycle	**Vincristine**	**Rituximab[a]**	**Cyclophosphamide**	**Doxorubicin**	**Prednisolone**
1	19%	0%	0%	0%	3%
2	19%	0%	1%	1%	4%
3	20%	0%	2%	1%	4%
4	19%	1%	2%	2%	4%
5	20%	-0%	2%	2%	3%
6	20%	0%	3%	2%	4%
7	20%	0%	3%	1%	5%
8	19%	0%	5%	1%	5%

Key: R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone.
[a] Negative values represent average doses that were higher than per protocol doses.

The distributions of second-line treatments for VR-CAP and R-CHOP (as incorporated in the economic model) are presented in Table 6.14 and based on the LYM-3002 trial.[9] All regimens received by 5% or more of patients in either treatment arm were included, except if they were labelled 'other' and the regimen could therefore not be identified. Patients receiving the infrequent treatments (i.e. received by <5%) were redistributed over the included regimens. For R-FC and R-bendamustine the distributions of second-line treatments were assumed to be equal as for R-CHOP.

Table 6.14: Distribution of second-line treatment by first-line treatment
(Retrieved from the model submitted by the company)

	CHOP	R-Bendamustine	R-Bortezomib	R-CHOP	Rituximab	Ibrutinib	Lenalidomide	Temsirolimus	Bendamustine
VR-CAP	7%	11%	2%	11%	13%	19%	13%	9%	15%
R-CHOP	12%	23%	10%	6%	18%	9%	10%	9%	4%

Key: CHOP, cyclophosphamide, doxorubicin, vincristine and prednisolone; IV, intravenous; R-bendamustine, rituximab with bendamustine; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; R-cytarabine, rituximab with cytarabine; R-FC, rituximab with fludarabine and cyclophosphamide; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone

Health state costs attributed to disease management

Health state costs attributed to disease management compromised of the costs related to tests, scans, medical visits, concomitant medication and transfusions (both red blood cell (RBC) and platelet transfusions). Concomitant medication and transfusion costs are only incorporated while on first-line treatment.

The resource use and costs related to tests, scans, medical visits (i.e. costs related to full blood count, biochemistry, blood glucose, computed tomography-scan, haematologist visit) are presented in table 6.15 (based on NHS reference costs).

Table 6.15: Medical resource use by health state
(based on Table 55 of the CS[2])

Item	**Unit price**	**Units per health state**			
		PFS (on treatment[a])	**PFS (off treatment[a])**	**At time of progression**	**Progressed**
Full blood count	£3.00	3 per treatment cycle	1 per 2-3 months[b]	1	0
Biochemistry	£1.18	3 per treatment cycle	1 per 2-3 months[b]	1	0
Blood glucose	£1.18	3 per treatment cycle	0	0	0
CT scan	£80.00	1 in treatment cycles 1, 3 and 6	0	1	0
Haematologist visit	£150.06	1 in treatment cycles 1, 3 and 6	1 per 2-3 months[b]	1	1 per 2-3 months[b]
Total costs per week		£43.70	£14.02	£234.24[c]	£13.64

Key: CT, Computed tomography
[a] First- or second-line
[b] This has been applied as once every 11 weeks
[c] Once only costs

Concomitant medication to treat AE or to avoid lowering the dose of chemotherapy were incorporated in the model. The percentage of patients receiving concomitant medication was retrieved from the LYM-3002 trial.[9] However, concomitant medications provided in the LYM-3002 trial[9] but unavailable in the UK were excluded from the analysis. The average concomitant medication costs were assumed for these patients. In addition, the percentage of patients receiving concomitant medication during treatment cycles 6, 7 and 8 was assumed to be the same (due to the low number of patients receiving concomitant medication in these treatment cycles).

Based on clinical advice, no additional administration visits were included for concomitant therapies to avoid double counting (i.e. it is assumed that concomitant medication will be administered during treatment administration or routine monitoring visits). Moreover, also based on clinical advice, it was assumed that patients self-administer G-CSFs (the impact of this assumption was explored in a scenario analysis).

For R-FC and R-bendamustine the concomitant medication costs were assumed to be equal as for R-CHOP. The company expected that this assumption would result in an overestimation of the concomitant medication costs for R-FC and R-bendamustine.

Unit prices of concomitant medication can be found in Table 57 of the CS[2]) while the percentage of patients receiving concomitant medication, the dosing and weekly costs of concomitant medication are presented in Table 6.16.

No concomitant medication costs were assumed for second-line treatment.

Table 6.16: Percentage of patients receiving concomitant medication, the dosing and weekly costs of concomitant medication (based on Table 56 of the CS[2])

	VR-CAP						R-CHOP						Dose per week (mg)	Cost per week
	Cycle						Cycle							
	1	2	3	4	5	6-8	1	2	3	4	5	6-8		
GCSF														
Filgrastim	19%	23%	25%	25%	28%	23%	16%	19%	19%	23%	23%	17%	0.46 mg	£97.06
Pegfilgrastim	6%	7%	8%	8%	9%	7%	5%	6%	6%	7%	8%	6%	2.00 mg	£228.79
Lenograstim	6%	7%	8%	8%	9%	7%	4%	4%	5%	5%	5%	4%	0.33 mg	£89.25
Antibacterials for systematic use														
Ciprofloxacin	13%	9%	8%	9%	7%	7%	14%	7%	7%	6%	4%	6%	7,000 mg	£0.38
Levofloxacin	13%	9%	8%	8%	7%	7%	6%	3%	3%	2%	2%	2%	3,500 mg	£1.67
Moxifloxacin	6%	4%	4%	4%	3%	3%	2%	1%	1%	1%	1%	1%	2,800 mg	£13.31
Ceftazidime	6%	4%	4%	4%	3%	3%	3%	1%	1%	1%	1%	1%	28,000 mg	£33.46
Ceftriaxone	6%	4%	4%	4%	3%	3%	3%	1%	1%	1%	1%	1%	14,000 mg	£7.14
Meropenem	4%	3%	2%	2%	2%	2%	3%	2%	2%	1%	1%	1%	21,000 mg	£75.01
Co-amoxiclav	10%	7%	7%	7%	5%	6%	7%	3%	3%	3%	2%	3%	13,125 mg	£2.84
Amoxicillin	3%	2%	2%	2%	2%	2%	3%	1%	1%	1%	1%	1%	10,500 mg	£0.48
Piperacillin + Tazobactam	6%	4%	4%	4%	3%	3%	4%	2%	2%	2%	1%	2%	126,000 mg	£36.62
Co-trimoxazole	16%	11%	10%	10%	8%	8%	12%	6%	6%	5%	4%	5%	13,440 mg	£2.99
Vancomycin	3%	2%	2%	2%	2%	2%	3%	1%	1%	1%	1%	1%	14,000 mg	£23.94
Amikacin	5%	3%	3%	3%	3%	3%	3%	1%	1%	1%	1%	1%	10,105 mg	£143.69
Aciclovir	38%	12%	13%	10%	10%	6%	27%	6%	4%	6%	4%	4%	5,600 mg	£0.87
Valaciclovir	7%	2%	3%	2%	2%	1%	3%	1%	1%	1%	0%	0%	21,000 mg	£7.04

Key: G-CSF, Granulocyte-colony stimulating agent; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone

Platelet transfusions are administered to avoid or treat thrombocytopenia, whereas RBC transfusions are used to avoid or treat anaemia. The number of transfusions for VR-CAP and R-CHOP were retrieved from the LYM-3002 trial (see Table 6.17).[9] Based on the finding that transfusions were predominantly administered between day 10 and 14 in the LYM-3002 trial,[9] it was assumed that, in the model, transfusions would be administered in week 2 of each treatment cycle. One IV administration visit was assumed for each transfusion. For R-FC and R-bendamustine the transfusion costs were assumed to be equal as for R-CHOP.

No transfusion costs were assumed for second-line treatment.

Table 6.17: Number of transfusions
(Based on Table 59 of the CS[2])

	Number of transfusions per patient per treatment cycle			
	RBC		Platelet	
Cycle number	VR-CAP	R-CHOP	VR-CAP	R-CHOP
Cycle 1	0.145	0.152	0.151	0.025
Cycle 2	0.121	0.137	0.126	0.023
Cycle 3	0.164	0.187	0.170	0.031
Cycle 4	0.102	0.128	0.106	0.021
Cycle 5	0.096	0.065	0.100	0.011
Cycle 6, 7 and 8	0.069	0.074	0.071	0.012
Cost per transfusion	£483.83[a]		£441.96[a]	

Key: R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; RBC; red blood cell; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone.
[a] This includes IV administration visit costs

The concomitant treatment and transfusion costs per treatment cycle (only incorporated during the first-line on-treatment phase) are presented in Table 6.18.

Table 6.18: Concomitant treatment and transfusion costs per treatment cycle
(Retrieved from the model submitted by the company)[a]

Cycle number	Concomitant medication	RBC transfusions	Platelet transfusions
	VR-CAP		
1	£163.42	£70.30	£66.69
2	£171.96	£58.74	£55.72
3	£178.92	£79.38	£75.30
4	£184.77	£49.48	£46.94
5	£192.58	£46.52	£44.13
6	£165.52	£33.20	£31.49
	R-CHOP		
1	£107.82	£73.60	£11.20
2	£100.46	£66.10	£10.06
3	£103.03	£90.45	£13.77
4	£116.32	£62.07	£9.45
5	£115.42	£31.55	£4.80
6	£91.16	£35.93	£5.47

Key: R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; RBC; red blood cell; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone.
[a] Please note that the costs reported in Table 60 of the CS[2] are incorrect (i.e. not consistent with the values incorporated in the economic model).

Adverse event costs

Treatment costs for the following adverse events were included in the model (during first-line treatment only):

- Febrile neutropenia Grade 3-5
- Pneumonia Grade 3-5
- Diarrhoea Grade 4/5
- Peripheral sensory neuropathy Grade 4/5
- Sepsis

To avoid double counting (with the costs related to transfusions, concomitant medication and/or dose reductions), treatment costs of the following adverse events were not incorporated in the model based on clinical advice:

- Neutropenia Grade 3-5
- Thrombocytopenia Grade 3-5
- Anaemia Grade 3-5
- Leukopenia Grade 3-5
- Lymphopenia Grade 3-5
- Fatigue Grade 3-5
- Diarrhoea Grade 3
- Peripheral sensory neuropathy Grade 2-3
- Alopecia

The AE costs per week are presented in Table 6.19. These weekly AE costs were calculated using the weekly probabilities reported in Table 6.7 and the unit costs (NHS reference costs) reported in Table 61 of the CS[2].

Table 6.19: AE costs per week
(Retrieved from the model submitted by the company)[a]

	VR-CAP	R-CHOP	R-FC	R-bendamustine
Febrile neutropenia	£20.90	£25.99	£5.34	£11.77
Pneumonia	£2.73	£1.72	£0.00	£0.81
Diarrhoea	£0.00	£0.15	£0.00	£0.11
Peripheral sensory neuropathy	£0.11	£0.00	£0.00	£0.00
Sepsis	£2.68	£0.95	£0.00	£0.00

Key: R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; R-bendamustine, rituximab with bendamustine; R-FC, rituximab with fludarabine and cyclophosphamide; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone.
[a] Please note that the costs reported in Table 62 of the CS[2] are incorrect (i.e. not consistent with the values incorporated in the economic model).

End of life costs

Based on a King's Fund report,[57] end-of-life costs were incorporated in the economic analysis. This report provided the average costs of community and acute care for patients with cancer in the last eight weeks of life. The reported £5,324 was inflated to the 2013/2014

price level leading to end-of-life costs of £6,018. The company mentioned that not all of these costs were direct NHS costs (some fell on 'third sector' healthcare organisations).

ERG comment: Regarding drug acquisition and administration costs, the mean second-line duration was used in the model as the maximum second-line duration, potentially underestimating the second-line treatment costs. However, this is likely to be conservative as patients spent least time in the second-line when treated with VR-CAP (see Table 6.24) compared with R-CHOP, R-FC and R-bendamustine.

The company calculated the drug costs using the per-protocol dosing and subsequently applied relative dose reductions for VR-CAP and R-CHOP only. For R-FC and R-bendamustine no dose reductions were assumed, which is not considered to be conservative. Moreover, it is questionable whether the dose reductions observed in the LYM-3002 trial[9] are representative for UK clinical practice, particularly because no UK patients were included in the LYM-3002 trial,[9] and only 28% of the patients are from the European Union. Additionally , the company highlighted that there might be differences between the trial data and UK clinical practice with regards to dose reduction in response to peripheral neuropathy (as in UK practice bortezomib is administered subcutaneously rather than intravenously possibly leading to a reduction in peripheral neuropathy, see response to clarification question C17[16]). Therefore, the ERG conservatively incorporated the per protocol dosing without dose reductions in the ERG base case. The response to clarification question C17 from the company indicated that this would increase the ICER for VR-CAP versus R-CHOP to £24,265.[16]

The health state costs include the costs related to tests, scans, medical visits, concomitant medication and transfusions (the latter two are only incorporated during first-line treatment). For the calculation of the concomitant medication costs the company mentioned that it was assumed that patients that received a concomitant medication in the LYM-3002 trial[9] that is unavailable in the UK would receive the average concomitant medication costs. However, this was not incorporated as such in the model submitted by the company. Also, there was an error in the calculation of the costs for pegfilgrastim for R-CHOP (the unit price for filgrastim was used). Both issues were adjusted in the ERG base case (the adjusted concomitant medication costs are reported in Table 6.20). Moreover, although this can be considered a reasonable assumption, it is unclear whether this is conservative or not. The ERG explored the impact of this assumption in an exploratory analysis assuming no concomitant and transfusion costs for R-FC and R-bendamustine. Moreover, the exclusion of concomitant medication and transfusions for second-line treatment is likely to be conservative, as more patients receive a second-line treatment after one of the comparator treatments in first-line.

Table 6.20: Concomitant treatment costs per treatment cycle (adjusted by ERG)

Cycle number	VR-CAP	R-CHOP
1	£88.05	£59.78
2	£93.45	£56.75
3	£97.36	£58.21
4	£100.57	£65.94
5	£105.05	£65.67
6	£90.20	£51.55
Key: R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; RBC; red blood cell; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone.		

Unit costs for the majority of adverse events were assumed to be zero to avoid double counting with concomitant medication and transfusion costs (for neutropenia, thrombocytopenia, anaemia, leukopenia, lymphopenia, fatigue, diarrhoea (grade 3 only), peripheral sensory neuropathy (only grade 2-3) and alopecia). Although, (part of) these costs were probably incorporated through including concomitant medication and transfusion costs (e.g. platelet costs are higher for VR-CAP might be due a higher occurrence of thrombocytopenia), the extent of this is unclear. Therefore, the ERG explored the impact of this assumption in an exploratory analysis by including AE costs for all AE except Lymphopenia (no costs were found for this AE). See Table 6.21 for the AE costs that were used in this exploratory analysis.

Table 6.21: Adverse event costs calculated by the ERG

Adverse event	ERG estimate	Source
Neutropenia Grade 3-5	£167.28	NHS reference costs 2013-2014; HRG code: XD25Z
Thrombocytopenia Grade 3	£570.97	NHS reference costs 2013-2014; NEI_S; weighted average of HRG codes: SA12G, H, J, and K
Thrombocytopenia Grade 4-5	£2,191.65	NHS reference costs 2013-2014; NEI_L; weighted average of HRG codes: SA12G, H, J, and K
Anaemia Grade 3	£516.66	NHS reference costs 2013-2014; NEI_S; weighted average of HRG codes: SA04G, H, J, K and L
Anaemia Grade 4-5	£1,853.10	NHS reference costs 2013-2014; NEI_L; weighted average of HRG codes: SA04G, H, J, K and L
Leukopenia Grade 3-5	£167.28	Costs assumed to be equal to neutropenia
Lymphopenia Grade 3-5	£0.00	No costs were found
Fatigue Grade 3-5	£12.00	NICE ERG report abiraterone (TA 259), table 24, p. 64.
Diarrhoea Grade 3	£572.80	NHS reference costs 2013-2014; NEI_S; HRG code: PF26B
Peripheral sensory neuropathy Grade 2/3	£417.33	NHS reference costs 2013-2014; NEI_S; HRG code: AA26H
Alopecia	£44.00	NICE ERG report rituximab (TA 243), table 41, p. 149.

The end-of-life costs included by the company were based on a small and heterogeneous patient sample (40 patients with cancer receiving palliative care). It is unclear whether these end-of-life costs are representative for the population considered in this assessment. Moreover, as mentioned by the company,[2] these end-of-life costs include costs for 'third

sector' (i.e. not-for-profit and non-governmental) healthcare organisations. Therefore, the ERG conservatively excluded these end-of-life costs in the ERG base case.

The ERG notes that the costs reported in Tables 60 and 62 of the CS[2] are incorrect (i.e. not consistent with the values incorporated in the economic model) and hence readers are recommended to consider Tables 6.18 and 6.19 of this report instead.

The ERG checked all unit prices provided by the company in Tables 48, 49, 55, 57, 59 and 61 of the CS.[2] The ERG could not reproduce the costs presented in Table 6.22. These costs were therefore corrected in the ERG base case.

Table 6.22: Unit prices reported in the CS that could not be reproduced by the ERG

Cost item	Company estimate	Table in CS[2]	ERG estimate	Source
Cyclophosphamide, oral (50 mg)	£0.82	48	£1.39	MIMS
Administration by district nurse	£65.00	49	£76.00[a]	PSSRU 2014
Haematologist visit	£150.06	55	£154.66	NHS reference costs 2013-2014[b]
Moxifloxacin (400 mg)	£1.90	57	£2.09	MIMS

[a] This estimate includes qualification costs (in contrast with the estimated costs used by the company).
[b] These costs are calculated by using a weighted average of HRG codes: WF01A and C (both clinical haematology)

6.2.10 Cost-effectiveness results

The company did provide most results for VR-CAP and R-CHOP only. However, for completeness and given that R-FC and R-bendamustine are also listed as comparators in the scope, the ERG summarised the cost-effectiveness results including all four comparators if these results could be retrieved from the economic model or the CS.[2]

Disaggregated cost and effects (deterministic)

VR-CAP resulted in the highest costs compared with the three comparators (see Table 6.23). These higher costs of VR-CAP could mainly be attributed to the difference in first-line treatment costs. This was reflected in the costs differences in the first line, pre-progression health state. In contrast, the second-line treatment costs were substantially lower for VR-CAP compared with the other three comparators.

Table 6.23: Disaggregated costs (discounted) per cost category and per health state
(Deterministic; retrieved from the model submitted by the company)

	VR-CAP	R-CHOP	Δ versus VR-CAP	R-FC	Δ versus VR-CAP	R-Benda	Δ versus VR-CAP
Per cost category							
First-line treatment	£22,606	£8,041	-£14,566	£10,246	-£12,361	£12,700	-£9,907
Administration	£5,817	£1,564	-£4,253	£1,559	-£4,258	£2,638	-£3,178
Adverse events	£458	£499	£41	£122	-£336	£270	-£188
Concomitant medication	£1,014	£606	-£408	£801	-£213	£765	-£249
Medical resource use[a]	£4,186	£4,671	£484	£4,550	£363	£4,515	£328
Second-line treatment	£7,152	£9,423	£2,271	£9,272	£2,120	£9,302	£2,150
Terminal care	£4,605	£4,821	£217	£4,821	£217	£4,821	£217
Total	***£45,838***	***£29,625***	***-£16,213***	***£31,370***	***-£14,467***	***£35,011***	***-£10,826***
Per health state							
First-line, pre-progression	£31,374	£11,942	-£19,433	£13,846	-£17,528	£17,454	-£13,921
First-line, post-progression	£290	£294	£4	£338	£48	£330	£41
Second-line, pre-progression	£7,729	£10,169	£2,440	£10,005	£2,277	£10,038	£2,310
Second-line, post-progression	£6,445	£7,221	£776	£7,181	£736	£7,189	£744
Total	***£45,838***	***£29,625***	***-£16,213***	***£31,370***	***-£14,467***	***£35,011***	***-£10,826***

Key: R-benda, rituximab with bendamustine; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; R-FC, rituximab with fludarabine, and cyclophosphamide; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone.
[a] Includes costs related to tests, scans, medical visits and transfusions (both red blood cell and platelet transfusions)

VR-CAP was the most effective treatment (see Table 6.24). Similar to the costs, the main gain in effect (both QALYs and LYs) of VR-CAP compared with the other treatments could be attributed to increased QALYs and LYs in the first-line pre progression health state. Moreover, the QALYs and LYs for VR-CAP were lower in the second-line health states compared with the other comparators

Table 6.24: Disaggregated QALYs and Lys (discounted) per health state
(Deterministic; retrieved from the model submitted by the company)

	VR-CAP	R-CHOP	Δ versus VR-CAP	R-FC	Δ versus VR-CAP	R-Benda	Δ versus VR-CAP
QALYs							
First-line, pre-progression	2.705	1.546	-1.159	1.546	-1.159	1.546	-1.159
AE Utility Decrement	-0.004	-0.003	0.001	0.000	0.003	-0.001	0.003
First-line, post-progression	0.137	0.103	-0.033	0.144	0.007	0.136	-0.001
Second-line, pre-progression	0.223	0.288	0.065	0.283	0.060	0.284	0.061
Second-line, post-progression	1.085	1.415	0.330	1.392	0.307	1.397	0.312
Total	***4.146***	***3.350***	***-0.796***	***3.364***	***-0.782***	***3.361***	***-0.785***
LYs							
First-line, pre-progression	3.541	2.023	-1.517	2.023	-1.517	2.023	-1.517
First-line, post-progression	0.197	0.149	-0.048	0.207	0.010	0.196	-0.001
Second-line, pre-progression	0.292	0.377	0.085	0.371	0.079	0.372	0.080
Second-line, post-progression	2.411	3.145	0.734	3.093	0.682	3.104	0.692
Total	***6.441***	***5.695***	***-0.746***	***5.695***	***-0.746***	***5.695***	***-0.746***

Key: LY, life year; QALY, quality-adjusted life year; R-benda, rituximab with bendamustine; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; R-FC, rituximab with fludarabine, and cyclophosphamide; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone.

Cost-effectiveness (probabilistic)

The probabilistic cost-effectiveness results are presented in Table 6.25. VR-CAP was most expensive and most effective whilst R-CHOP was least effective and least expensive. The ICERs compared with VR-CAP ranged between £13,725 (versus R-bendamustine) and £20,264 (versus R-CHOP).

Table 6.25: Cost-effectiveness results
(Probabilistic; retrieved from the model submitted by the company)

Treatment	Expected outcomes		Compared with VR-CAP		
	Costs	QALY	ΔCosts	ΔQALY	ICER
VR-CAP	£45,724	4.142			
R-CHOP	£29,491	3.341	£16,234	0.801	£20,264
R-FC	£31,222	3.355	£14,503	0.787	£18,430
R-bendamustine	£34,883	3.352	£10,842	0.790	£13,725

Key: ICER, incremental cost-effectiveness ratio; R-bendamustine, rituximab with bendamustine; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; R-FC, rituximab with fludarabine, and cyclophosphamide; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone.

The full incremental analysis done by the ERG indicated that R-bendamustine is dominated by R-FC, and that R-FC is extendedly dominated by VR-CAP. Hence, the only relevant ICER remaining is that of VR-CAP versus R-CHOP (Table 6.26).

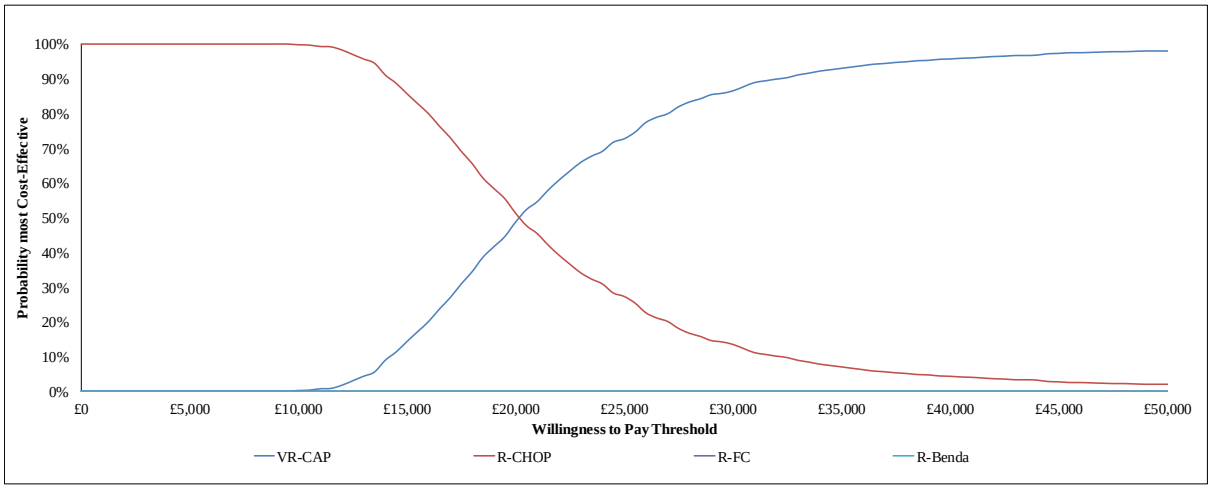
Table 6.26: Full incremental analysis of cost-effectiveness results
(Probabilistic; retrieved from the model submitted by the company)

Treatment	Expected outcomes		Incremental analysis			
	Costs	QALYs	Comparator	Costs	QALYs	ICER
R-CHOP	£29,491	3.341				
R-FC	£31,222	3.355	R-CHOP	£1,731	0.014	Extended dominance
R-bendamustine	£34,883	3.352	R-CHOP	£5,392	0.011	Dominance
VR-CAP	£45,724	4.142	R-CHOP	£16,234	0.801	£20,264

Key: ICER, incremental cost-effectiveness ratio; R-benda, rituximab with bendamustine; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; R-FC, rituximab with fludarabine, and cyclophosphamide; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone.

The cost-effectiveness acceptability curves (Figure 6.7) indicate that, at a willingness to pay (WTP) threshold of £20,000, R-CHOP has the highest probability of being cost-effective (51.3%) followed by VR-CAP (48.7%). The probabilities of being cost-effective for R-FC and R-bendamustine were 0.0%. VR-CAP has the highest probability (86.5%) of being cost-effective at a WTP threshold of £30,000, followed by R-CHOP (13.5%), R-FC (0.0%) and R-bendamustine (0.0%).

Figure 6.7: Cost-effectiveness acceptability curves
(Retrieved from the model submitted by the company)



Key: R-benda, rituximab with bendamustine; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; R-FC, rituximab with fludarabine, and cyclophosphamide; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone.

ERG comment: Despite being listed in the scope, the company did not consider R-FC and R- bendamustine in the base case results. The ERG did incorporate these comparators in the description of the cost-effectiveness results to be consistent with the scope. Nevertheless, in the base case of the company the cost-effectiveness results show that R-FC and R-bendamustine are (extendedly) dominated, which implies that for decision making only the comparison of VR-CAP with R-CHOP is relevant (probabilistic ICER: £20,264).

Not all potential stochastic parameters were incorporated as stochastic parameters in the economic model. As a result the parameters presented in Tables 53 (percentage of patients receiving first-line treatment) and 56 (percentage of patients receiving concomitant medication) of the CS[2] are not considered in the deterministic sensitivity analysis (DSA) and the uncertainty is not reflected in the probabilistic sensitivity analysis (PSA). The company performed the PSA while including the parameters from these Tables stochastically in response to clarification question C19.[16] This analysis increased the ICER between VR-CAP and R-CHOP to £20,642. Also, the probabilities for VR-CAP being cost-effective decreased; these are 44.6% and 87.9% for WTP thresholds of £20,000 and £30,000 respectively. The ERG prefers to include all potential stochastic parameters as stochastic in the PSA. Hence this was adjusted in the ERG base case. In addition, the company also included unit prices, age and weight as stochastic parameters in the PSA. The ERG would not incorporate these parameters in the PSA as the uncertainty in these parameters do not reflect parameter uncertainty. Therefore, unit prices, age and weight were not included as stochastic parameters in the ERG base case.

6.2.11 Sensitivity analyses

Deterministic sensitivity analysis

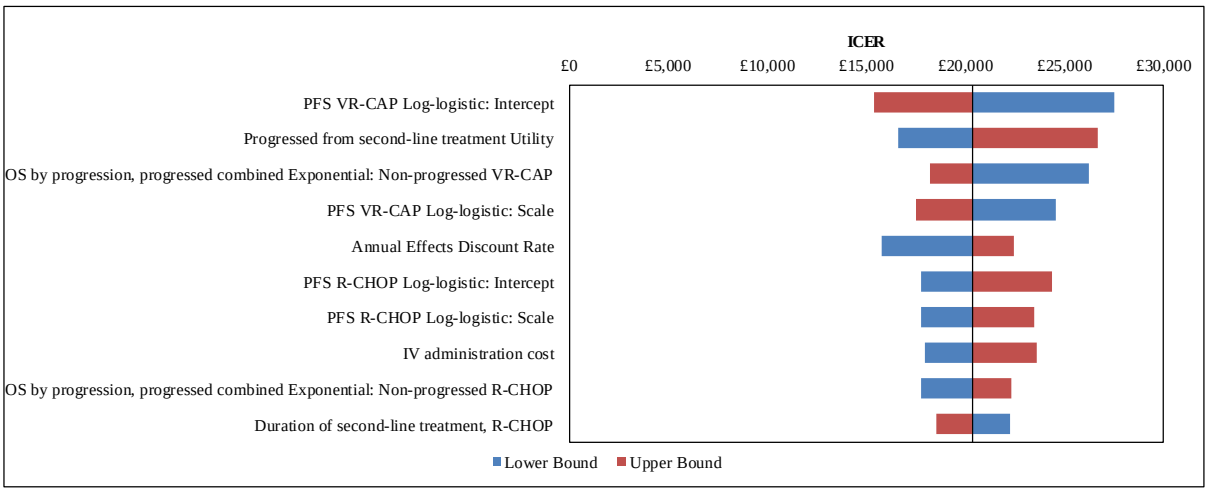
Deterministic sensitivity analyses were performed for parameters that were included stochastically in the economic model and the annual discount rates. Stochastic parameters were varied within the 95% confidence interval of the distribution that was assigned and discount rates were varied between 0% and 5%. The 10 most influential parameters are presented in tornado diagrams (Figure 6.8).

Considering the stochastic parameters, the ICERs were most sensitive to the survival functions used to model PFS and OS, the utility for patients progressed from second-line treatment, IV administration costs and the duration of second-line treatment.

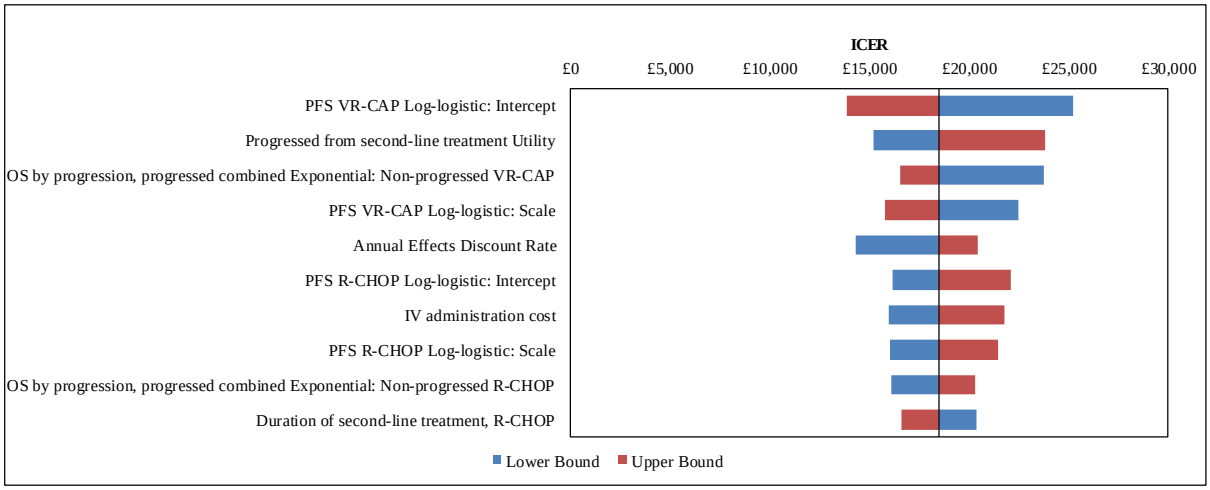
Figure 6.8: Tornado diagram displaying the ICER sensitivity of the 10 most influential model inputs

(Deterministic; retrieved from the model submitted by the company)

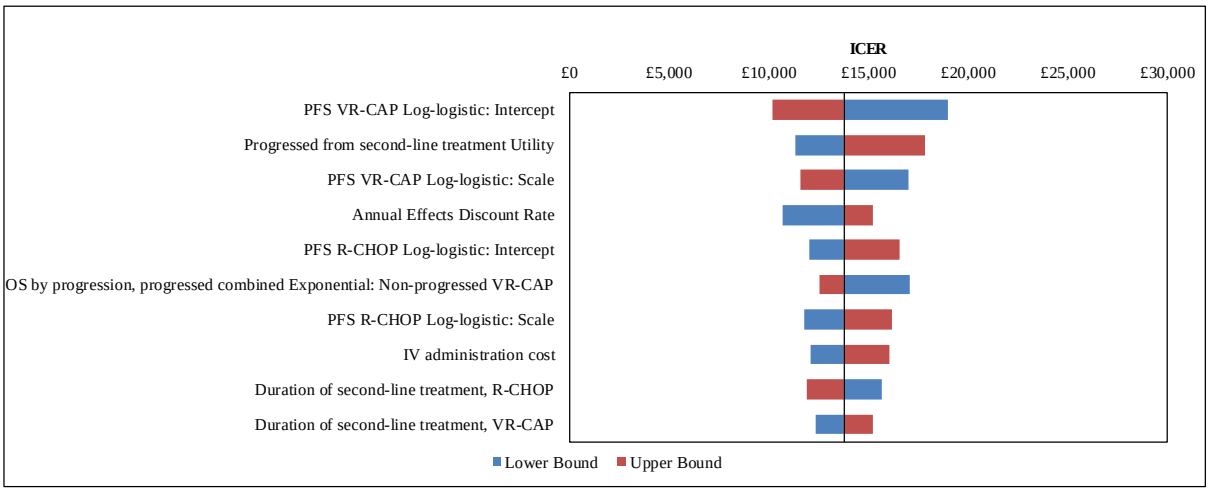
A: VR-CAP versus R-CHOP



B: VR-CAP versus R-FC



C: VR-CAP versus R-bendamustine



Key: R- bendamustine, rituximab with bendamustine; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; R-FC, rituximab with fludarabine, and cyclophosphamide; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone.

Deterministic scenario analysis

The company performed a large number of scenario analyses for the comparison between VR-CAP and R-CHOP (methods are described in Tables 63 and 64 of the CS.[2]) The deterministic results of these analyses were reported in the Tables 71 and 72 of the CS.[2]

The most influential scenario analyses where those incorporating different parametric distributions for PFS; using Weibull, Gamma and Gompertz distributions increased the ICER to £25,849, £27,697 and £30,452 respectively. Moreover, changing the utility value for patients progressed from second-line treatment to 0.693 (equal as for patients progressed from first-line treatment) increased the ICER to £26,241. Changing all health state utility values to correspond with those from Doorduijn et al[41] (i.e. 0.61 for progression free in the first- and second-line and 0.45 for progressed patients in the first- and second-line) did increase the ICER to £28,746. Furthermore, when assuming no second-line treatment (i.e. patients were assumed to stay in the first-line post-progression health state; this did not impact survival), this would increase the ICER to £29,858. In the two scenario analyses including R-maintenance (using HRs (PFS; OS) of 0.55; 1.00 and 0.41; 0.48) resulted in ICERs of £22,518 and £23,040 respectively. When assuming a time horizon of 10 year the ICER would increase to £27,443. Finally, when setting the baseline age to 75 (base case value = 69 years) the ICER increased to £26,010.

One additional scenario analysis was presented in table 75 of the CS[2] which included all four comparators. In this scenario the assumption of equal efficacy between R-CHOP, R-FC and R-bendamustine was relaxed and the results of the Bucher indirect comparison were used (indirect comparison is presented in table 24 of the CS[2]). Note that PFS for R-FC was still assumed to be equal to R-CHOP (as R-FC was not included in the PFS indirect comparison). The full incremental analysis of this scenario (table 6.27) indicated that R- bendamustine would be cost-effective for WTP thresholds between £3,299 and £30,640, while VR-CAP is cost-effective for WTP thresholds above £30,640.

Table 6.27: Scenario analysis while incorporating indirect comparison for PFS and OS (Deterministic; retrieved from the model submitted by the company)

Treatment	Expected outcomes		Incremental analysis			
	Costs	QALYs	Comparator	Costs	QALYs	ICER
R-FC	£29,616	2.584				
R-CHOP	£29,625	3.350	R-FC	£9	0.766	£11
R-bendamustine	£30,612	3.649	R-CHOP	£987	0.299	£3,299
VR-CAP	£45,838	4.146	R-Benda	£15,226	0.497	£30,640

Key: ICER, incremental cost-effectiveness ratio; R-benda, rituximab with bendamustine; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; R-FC, rituximab with fludarabine, and cyclophosphamide; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone.

Threshold analysis

The company performed threshold analyses examining the OS hazard ratio (HR) needed for R-FC versus R-CHOP and R-bendamustine versus R-CHOP in order for these treatments to become cost-effective at WTP thresholds of £20,000 and £30,000. Note that, except for OS, equal efficacy between R-CHOP, R-FC and R-bendamustine was still assumed in this

analysis. This threshold analysis showed that at a WTP threshold of £20,000 the HRs should be at least 0.95 for R-FC versus R-CHOP and at least 0.83 for R-bendamustine versus R-CHOP in order for R-FC and R-bendamustine to be cost-effective. At a WTP threshold of £30,000 these HRs were 0.79 and 0.73 respectively.

Finally, the company performed a threshold analysis to examine the potential impact of R-maintenance therapy. It was examined what HRs were required for R-maintenance compared to no R-maintenance therapy in order for VR-CAP to become cost-effective at WTP thresholds of £20,000 and £30,000. The results of this analysis are presented in Table 78 of the CS.[2]

ERG comment: Although a large number of deterministic sensitivity analyses was performed, in the opinion of the ERG the DSA for the parameters presented in Tables 53 (percentage of patients receiving first-line treatment) and 56 (percentage of patients receiving concomitant medication) of the CS[2] were missing (as described earlier). The company performed these DSAs in response to clarification question C19[16] for VR-CAP versus R-CHOP only. This did not change the tornado diagram with the 10 most influential model inputs (presented in Figure 6.8). No updated tornado diagrams were provided for VR-CAP versus R-FC and VR-CAP versus R-bendamustine.

The company did not consider R-FC and R-bendamustine in the scenario analyses. Therefore, the impact of the scenario analyses on the comparisons with R-FC and R-bendamustine are unclear.

6.2.12 Subgroup analysis

No subgroup analysis was performed by the company.

ERG comment: Although no subgroups of interest were listed in the scope, Figure 8 in the CS did highlight a potential reduction in relative effectiveness of VR-CAP versus R-CHOP regarding PFS for patients from the European Union/North American region.[2] This was confirmed for patients from the European Union by Table 2 in the response to clarification question B2.[16] Although requested in clarification question C4, the company did not provide a subgroup analysis for the European Union or European Union/North American region subgroup. As the treatment effectiveness seems lower for the European Union subgroup, the relative treatment effect for PFS was conservatively adjusted to reflect the European Union subgroup in the ERG base case.

6.2.13 Model validation and face validity check

Face validity

The conceptual model and key assumptions were validated using an advisory board with practising UK haematologists and two UK health economic experts.[7] These individuals were selected as leading health economic and methodology development experts in the UK, one of whom is a regular member of a NICE Evidence Review Group, and the other is a regular Scottish Medicines Consortium reviewer. The model was presented to the health economists during the advisory board. The shared content comprised the clinical trial data package, the model structure, the assumptions regarding the treatment pathway, the survival analysis undertaken, the preliminary EQ-5D analysis undertaken and the methodology used for costs

and resource use. The opinions provided by these external health economists were used to determine the model base case in terms of survival analysis and utilities.

Internal validity

An economist not involved in the model adaptation reviewed the model for coding errors, inconsistencies and the plausibility of inputs. The model was also put through an internal and external checklist of known modelling errors and the assumptions were questioned.[58]

Cross-validation

No cross-validation was performed, presumably because the systematic review of cost-effectiveness studies did not identify any relevant cost-effectiveness studies.

External validity

Long-term trial projections compared well to the available Kaplan–Meier data (see CS[2] Figures 33 and 34). The modelled median PFS felt within the confidence interval of the median PFS observed from the trial. This was not assessable for the OS (see CS[2] Table 66). Survival was lower for both arms of the model than in the general population (as expected).

ERG comment: It appears that the appropriate validity checks were undertaken by the company and the results are satisfactory. Re-running the model, the base case deterministic and probabilistic results by the ERG confirm the reported findings from the company. The extreme value analyses undertaken by the ERG resulted in outcomes in the expected direction. However, as indicated in earlier 'ERG comment' sections, some programming errors were detected, making a corrected ERG base case necessary.

6.3 Exploratory and sensitivity analyses undertaken by the ERG

6.3.1 New ERG base case analysis

Based on several remarks in Section 6.2 of this report the ERG defined a new base case analysis. This new ERG base case included the following corrections and adjustments:

Correction of errors in the model

- Correct the unit prices which were different in the reference price list
- Correct error in the calculation of adverse events
- Correct calculation of costs concomitant medication
- Include half-cycle correction
- Age, weight and unit prices were made fixed instead of stochastic
- Proportion of patients receiving treatment during a cycle and proportion of patients receiving concomitant medication were made stochastic to reflect second order uncertainty

Adjustments to the model

- Adjust PFS according to the HR of the EU population
- Start second line treatment at time of progression
- Utility for progression from second line treatment is calculated (0.624) by subtracting the disutility as found in Doorduijn et al[41] from the utility in patients with PFS from first line treatment

- Exclude end of life costs
- Use per protocol dosage instead of observed dose reductions since it is unknown whether the dose reduction are applicable to UK patients
- The primary assessment of progression is used instead of the alternative assessment
- Indirect treatment comparison is used for the effectiveness of R-FC and R-bendamustine instead of assuming equal effectiveness as R-CHOP
- Overall survival is not differentiated between patients with and without progression, but between treatments instead
- Exclude all-cause mortality as this is already incorporated in the overall survival estimate
- The exponential distribution is used for the extrapolation of PFS in the VR-CAP arm and the log-logistic distribution is used for the extrapolation of PFS in the R-CHOP arm

The results of the adjusted ERG base case are presented in Table 6.28. It can be seen that in the new base case, the ICER of VR-CAP versus R-CHOP has increased by £14,000 to £34,039 compared to the CS base case. The ICER of VR-CAP versus R-FC changed from £18,430 to £13,455 and the ICER of VR-CAP versus R-bendamustine increased from £13,725 to £16,762. A full incremental analysis is shown in Table 6.28 and shows that R-FC and R-bendamustine are dominated by R-CHOP, i.e. they have higher costs and lower QALYs. The difference between the results of the company and the ERG base case are mainly caused by the following changes in the model:

- The use of a higher utility value for patients with progression from second line treatment. This effect is most evident in the R-CHOP treatment arm as the proportion of patients in this health state is the largest compared to the other three treatments.
- The use of per-protocol dosage instead of observed dose reductions
- The use of the exponential distribution to extrapolate the PFS of the VR-CAP arm since more patients progress according to the exponential distribution as compared to the log-logistic distribution
- The estimation of overall survival for all patients, while excluding all-cause mortality
- Treatment effectiveness of R-FC and R-bendamustine based upon indirect treatment comparison instead of the assumption of similar efficacy to that of R-CHOP.

Table 6.28: New ERG base case – pairwise comparison to VR-CAP

Treatment	Expected outcomes		Compared with VR-CAP		
	Costs	QALY	ΔCosts	ΔQALY	ICER
VR-CAP	£44,140	4.57			
R-CHOP	£22,186	3.93	£21,955	0.64	£34,039
R-FC	£22,370	2.96	£21,770	1.62	£13,455
R-bendamustine	£24,957	3.43	£19.184	1.14	£16,762

Key: ICER, incremental cost-effectiveness ratio; R-bendamustine, rituximab with bendamustine; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; R-FC, rituximab with fludarabine, and cyclophosphamide; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone.

Table 6.29: New ERG base case – full incremental comparison

Treatment	Expected outcomes		Incremental analysis			
	Costs	QALY	Comparator	ΔCosts	ΔQALY	ICER
R-CHOP	£22,186	3.93				
R-FC	£22,370	2.96	R-CHOP	£184	-0.97	Dominated by R-CHOP
R-bendamustine	£24,957	3.43	R-CHOP	£2,771	-0.5	Dominated by R-CHOP
VR-CAP	£44,140	4.57	R-CHOP	£21,955	0.64	£34,039

Key: ICER, incremental cost-effectiveness ratio; R-bendamustine, rituximab with bendamustine; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; R-FC, rituximab with fludarabine, and cyclophosphamide; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone.

Probabilistic sensitivity analysis

A PSA was performed for all comparators to capture the uncertainty in the estimation of input parameters in the new ERG base case. Figure 6.9 presents the cost-effectiveness plane and Figure 6.10 shows the cost-effectiveness acceptability curves (CEACs) with all comparators. The probability that VR-CAP is cost-effective at a threshold of £20,000 and £30,000 is much smaller in the ERG base-case compared to the company's base-case (11% versus 49% and 39% versus 89% for a threshold of £20,000 and £30,000, respectively). Similar to the company's base case, the probability that R-FC or R-bendamustine are cost-effective at the usual NICE thresholds are negligible.

Figure 6.9: Cost-effectiveness plane for all treatment options (QALYs)

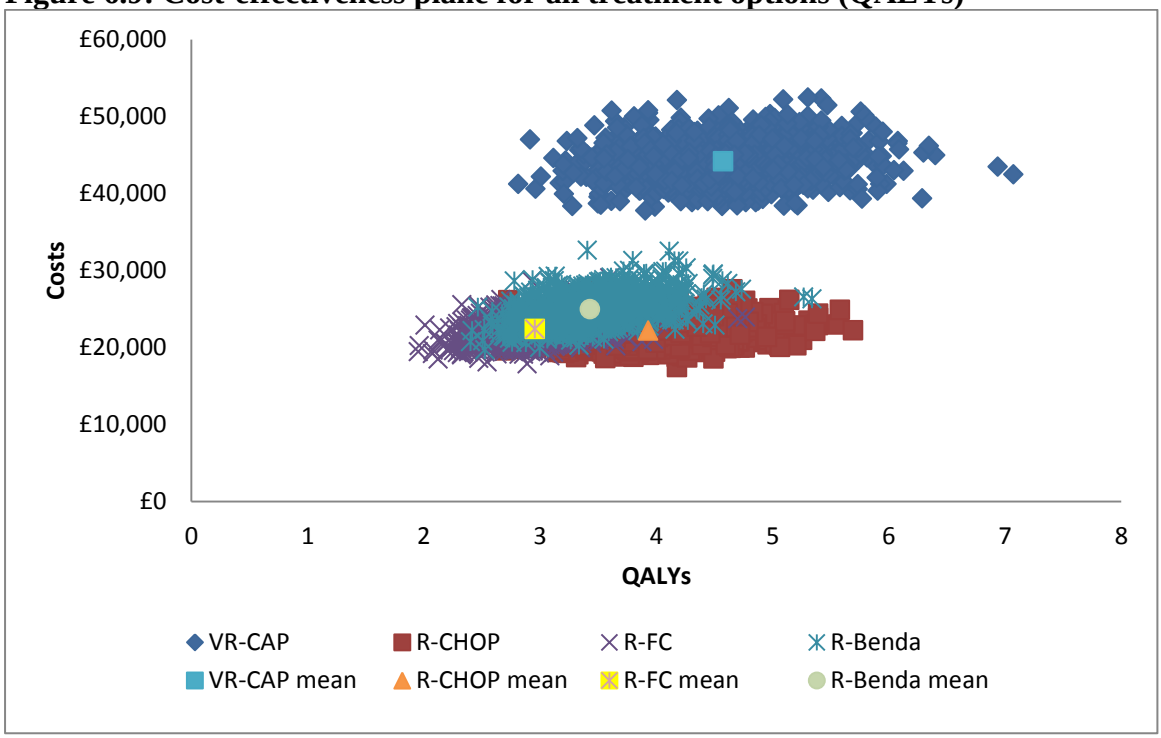
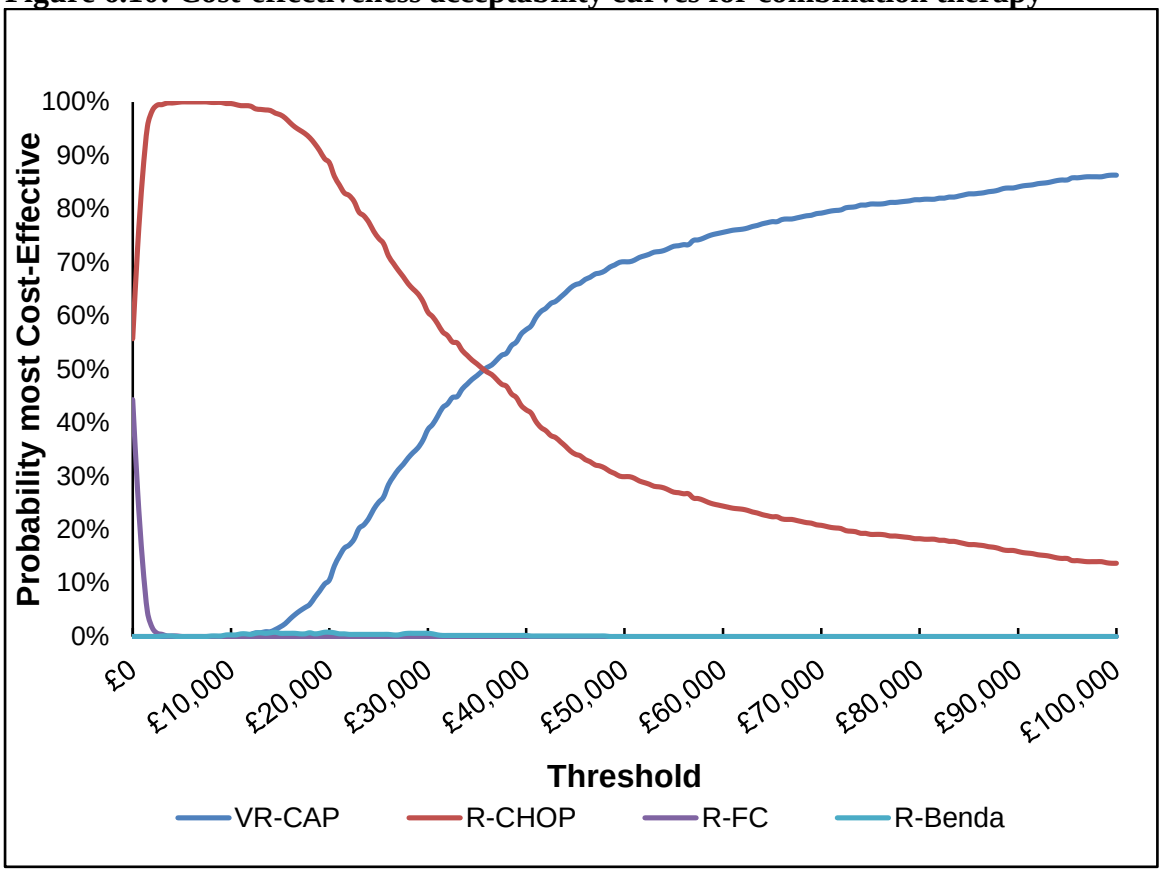


Figure 6.10: Cost-effectiveness acceptability curves for combination therapy



6.3.2 Exploratory scenario analyses undertaken by the ERG

1. *Exclude sepsis and alopecia as adverse events*
   The company stated that adverse events with a grade ≥3 and incidence above 5% in the VR-CAP or R-CHOP arm in the LYM-3002 trial. From the clinical study report it is clear that alopecia and sepsis did not fulfil this criterion; they were included based on clinical expert advice. To assess the impact of this, the ERG performed an additional analysis in which sepsis and alopecia were excluded.

2. *Include costs of adverse events for all adverse events included in the model*
   For some adverse events, the costs were not included by the company as these were assumed to be already incorporated in the costs of transfusions and concomitant medication. This assumptions neglects additional hospital days and visits due to the adverse events. Therefore, an additional analysis was performed including costs of all adverse events.

3. *Include costs of adverse events for all adverse events and exclude transfusion costs and costs of concomitant medication*
   There is a risk of double counting if both the costs of adverse events and the costs of transfusions and concomitant medications were included. Therefore, an additional analysis was performed in which costs of adverse events were included, but the costs of transfusions and concomitant medications were excluded.

4. *No transfusion and medication costs for R-FC and R-bendamustine*
   Since information regarding transfusions and concomitant medications are not available for R-FC and R-bendamustine, the company assumed that the costs of the transfusions and concomitant medications were equal to those for R-CHOP. This might be overestimation of the costs as the adverse event rates were smaller for R-FC and R-bendamustine. Therefore, an additional analysis was performed in which the costs of transfusions and concomitant medication were set to 0. This can be interpreted as the worst case scenario regarding these costs in a comparison with VR-CAP.

5. *Disutility for progression after second-line is only based upon the disutility for patients with low risk*
   In the current ERG base case, the disutility for progression after second-line treatment is calculated as the average disutility (0.14) for patients with low and high risk as reported by Doorduijn et al. It can be assumed that the disutility for patients with high risk is relatively small as for these patients the baseline utility is already quite low (0.44). Since the baseline utility in the LYM-3002 trial was much higher (0.762), an additional analysis was performed in which the disutility was only derived from the low risk group (0.24).

6. *Quality of life for patients with PFS from second-line treatment is worse than quality of life for patients with PFS from first line treatment*
   Since quality of life was not measured after progression from first-line treatment, it was assumed in the company base-case that patients with PFS from second-line treatment have a similar quality of life as patients with PFS from first-line treatment. However, it can be argued that the quality of life is slightly worse since disease has recurred after fist-line treatment and fewer treatment options are available to cure the disease. Therefore, an additional analysis has been performed in which the quality of life utility of PFS from second-line treatment was assumed to be 0.715. This utility was measured in the LYM-3002 trial in patients with progression from first-line treatment (based upon the primary assessment of progression).

7. *Non-stratified model for PFS*
   The motivation for the use of a stratified model is debatable based upon Figure 19 of the CS. The log-log curve only deviates at the end of follow-up. However, at that time, the number of patients at risk is much smaller. Therefore, an additional analysis has been performed in which a non-stratified model for PFS was used.

8. *Assume equal survival for both treatment arms*
   Due to the relatively immature survival data, uncertainty exists regarding the efficacy of VR-CAP in terms of overall survival. Therefore, as additional analysis was performed in which it was assumed that overall survival was similar between VR-CAP and R-CHOP

9. *Include R maintenance*
   Although it has been stated in a NICE scoping meeting that R-maintenance should be administered after any standard induction treatment, it is unclear whether it is actually administered in the UK. Furthermore, the effect of R-maintenance on OS and PFS is unclear. Therefore, two additional analyses have been performed using the HR of two different studies (similar to the company submission).
10. *Include a treatment-free interval after progression*
    The treatment-free interval that was included by the company was incorrect as patients without progression were also included in the estimation. Since no estimate was available of the treatment-free interval in patients *with* progression, the ERG performed two additional exploratory analyses regarding the length of the treatment-free interval. The first analysis assumes a post-progression treatment-free interval of one year for both treatments, whilst the second analysis assumes a post-progression treatment-free interval of 0.5 year after R-CHOP and one year after VR-CAP.

The results of all analyses are shown in Tables 6.30 to 6.32.

For the comparison of VR-CAP versus R-CHOP, one analysis has a very large impact, increasing the ICER to £328,757. This ICER is achieved when the OS is assumed to be the same for VR-CAP and R-CHOP. This assumption yields a QALY gain that is a factor 10 smaller than for the ERG base case, leading to a 10-fold larger ICER.

For all other analyses, the ICERs remained quite close to the base case, ranging from £32,518 (if an unstratified model is assumed for PFS) to £43,779 (if R-maintenance is included).

For the comparisons of VR-CAP versus R-FC and R-bendamustine we observe a similar range of outcomes (£12,911 to £23,478 and £16,470 to £29,520, respectively), though here we do not see the extreme change in ICER when the OS is assumed to be the same for VR-CAP and R-CHOP.

Table 6.30: Results exploratory analyses ERG; pairwise comparison VR-CAP and R-CHOP

	VR-CAP		R-CHOP		Incremental		ICER
	QALYs	Costs	QALYs	Costs	QALYs	Costs	Cost per QALY gained
ERG-base case	**4.57**	**£44,368**	**3.94**	**£22,467**	**0.63**	**£21,900**	**£34,738**
Exclude sepsis & alopecia	4.57	£44,278	3.94	£22,451	0.63	£21,826	£34,659
Include additional costs adverse events	4.57	£45,552	3.94	£22,847	0.63	£22,706	£36,015
Include additional costs adverse events, exclude costs concomitant medication and transfusion	4.57	£43,315	3.94	£21,455	0.63	£21,860	£34,674
No transfusion and concomitant medication costs for R-FC and R-benda	4.57	£44,368	3.94	£22,467	0.63	£21,900	£34,738
Disutility Doorduijn is only based on low risk group	4.20	£44,368	3.61	£22,467	0.59	£21,900	£37,202
Quality of life in PFS from second line is similar to the quality of life in patients with progression from first line treatment	4.56	£44,368	3.92	£22,467	0.63	£21,900	£34,728
Unstratified model for PFS	4.64	£42,828	3.98	£21,485	0.66	£21,343	£32,518
Survival VR-CAP is similar to survival R-CHOP	4.29	£43,591	4.22	£23,166	0.06	£20,425	£328,757
An average treatment-free interval of 365 days for all treatments	4.62	£43,238	3.99	£21,089	0.63	£22,149	£35,103
An average treatment-free interval of 365 days for VR-CAP and 182.5 for the other treatments	4.62	£43,238	3.97	£21,760	0.66	£21,477	£32,756
Include R-maintenance (source HR from NICE TA226)	4.66	£56,623	4.05	£31,338	0.61	£25,285	£41,585
Include R-maintenance (source HR from Kluin-Nelemans)	5.18	£58,765	4.60	£33,385	0.58	£25,379	£43,779

Key: HR, hazard ratio; ICER, incremental cost-effectiveness ratio; PFS, progression-free survival; R-bendamustine, rituximab with bendamustine; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; R-FC, rituximab with fludarabine, and cyclophosphamide; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone.

Table 6.31: Results exploratory analyses ERG; pairwise comparison VR-CAP and R-FC

	VR-CAP		R-FC		Incremental		ICER
	QALYs	Costs	QALYs	Costs	QALYs	Costs	Cost per QALY gained
ERG-base case	**4.57**	**£44,368**	**2.94**	**£22,568**	**1.63**	**£21,800**	**£13,378**
Exclude sepsis & alopecia	4.57	£44,278	2.94	£22,568	1.63	£21,710	£13,320
Include additional costs adverse events	4.57	£45,552	2.94	£22,830	1.63	£22,722	£13,943
Include additional costs adverse events, exclude costs concomitant medication and transfusion	4.57	£43,315	2.94	£21,112	1.63	£22,203	£13,625
No transfusion and concomitant medication costs for R-FC and R-benda	4.57	£44,368	2.94	£19,668	1.63	£24,699	£15,157
Disutility Doorduijn is only based on low risk group	4.20	£44,368	2.76	£22,568	1.44	£21,800	£15,121
Quality of life in PFS from second line is similar to the quality of life in patients with progression from first line treatment	4.56	£44,368	2.93	£22,568	1.63	£21,800	£13,402
Unstratified model for PFS	4.64	£42,828	2.97	£21,340	1.66	£21,488	£12,911
Survival VR-CAP is similar to survival R-CHOP	4.29	£43,591	2.71	£21,540	1.58	£22,051	£13,948
An average treatment-free interval of 365 days for all treatments	4.62	£43,238	2.98	£21,035	1.64	£22,203	£13,512
An average treatment-free interval of 365 days for VR-CAP and 182.5 for the other treatments	4.62	£43,238	2.96	£21,781	1.66	£21,457	£12,914
Include R-maintenance (source HR from NICE TA226)	4.66	£56,623	3.05	£18,859	1.62	£37,764	£23,355
Include R-maintenance (source HR from Kluin-Nelemans)	5.18	£58,765	3.54	£20,372	1.64	£38,393	£23,478

Key: HR, hazard ratio; ICER, incremental cost-effectiveness ratio; PFS, progression-free survival; R-bendamustine, rituximab with bendamustine; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; R-FC, rituximab with fludarabine, and cyclophosphamide; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone.

Table 6.32: Results exploratory analyses ERG; pairwise comparison VR-CAP and R-bendamustine

	VR-CAP		R-Bendamustine		Incremental		ICER
	QALYs	Costs	QALYs	Costs	QALYs	Costs	Cost per QALY gained
ERG-base case	**4.57**	**£44,368**	**3.41**	**£25,071**	**1.16**	**£19,297**	**£16,689**
Exclude sepsis & alopecia	4.57	£44,278	3.41	£25,071	1.16	£19,207	£16,610
Include additional costs adverse events	4.57	£45,552	3.41	£25,091	1.16	£20,461	£17,696
Include additional costs adverse events, exclude costs concomitant medication and transfusion	4.57	£43,315	3.41	£23,392	1.16	£19,923	£17,231
No transfusion and concomitant medication costs for R-FC and R-benda	4.57	£44,368	3.41	£22,248	1.16	£22,120	£19,131
Disutility Doorduijn is only based on low risk group	4.20	£44,368	3.23	£25,071	0.97	£19,297	£19,903
Quality of life in PFS from second line is similar to the quality of life in patients with progression from first line treatment	4.56	£44,368	3.40	£25,071	1.15	£19,297	£16,753
Unstratified model for PFS	4.64	£42,828	3.46	£23,309	1.17	£19,518	£16,673
Survival VR-CAP is similar to survival R-CHOP	4.29	£43,591	3.16	£23,818	1.13	£19,773	£17,568
An average treatment-free interval of 365 days for all treatments	4.62	£43,238	3.45	£23,909	1.17	£19,329	£16,470
An average treatment-free interval of 365 days for VR-CAP and 182.5 for the other treatments	4.62	£43,238	3.43	£24,475	1.19	£18,763	£15,766
Include R-maintenance (source HR from NICE TA226)	4.66	£56,623	3.50	£22,401	1.16	£34,222	£29,457
Include R-maintenance (source HR from Kluin-Nelemans)	5.18	£58,765	4.02	£24,481	1.16	£34,284	£29,520

Key: HR, hazard ratio; ICER, incremental cost-effectiveness ratio; PFS, progression-free survival; R-bendamustine, rituximab with bendamustine; R-CHOP, rituximab with cyclophosphamide, doxorubicin, vincristine and prednisolone; R-FC, rituximab with fludarabine, and cyclophosphamide; VR-CAP, bortezomib with rituximab, cyclophosphamide, doxorubicin and prednisolone.

6.4 Conclusions of the cost-effectiveness section

The economic model described in the CS is considered by the ERG to meet the NICE reference case to a reasonable extent. Reviewing the overall evidence, the ERG confirmed that there was no existing cost-effectiveness model for bortezomib for the current indication.

In terms of population, there is uncertainty regarding the generalisability of the LYM-3002 trial population to the population seen in clinical practice in the UK.

In terms of comparators, the CS did not include R-bendamustine and R-FC in the base case analysis, but presented the results of these treatments as 'exploratory' only, despite the fact that these two comparators were mentioned in the scope. The company justified this decision by explaining that these comparators are indicated for patients with frailty due to advanced age and/or comorbidities. Thus, R-bendamustine and R-FC would not be used for the same indication as VR-CAP.

The ERG thinks that exclusion of R-FC and R-bendamustine in the base case is inconsistent with the scope. Therefore, the ERG included these comparators in all analyses, not only as 'exploratory comparators'. In addition, the ERG thinks that the exclusion of post induction R-maintenance treatment is debatable since at the scoping meeting clinical experts indicated that maintenance treatment is common practice in the UK.

The cost-effectiveness results were generally robust under the sensitivity and scenario analyses conducted, with no scenarios bringing the ICER of VR-CAP versus R-CHOP above £30,000. The ICER was most sensitive to variations within the fit of the parametric models for PFS and OS. The model outcomes depend heavily on the PFS and OS curves used and the shape these curves have. The utility of patients who have progressed from second-line treatment was the parameter with the second greatest ICER sensitivity, due to patients spending a relatively long time in this health state.

The company model follows a logical structural with respect to the nature of the disease. A limitation of the model is that the additional treatments after second-line treatment are excluded. However, the ERG considers this as conservative. In contrast, the ERG did not agree with the company to not apply half cycle correction and to use the ITT population of the LYM-3002 trial to assess the effectiveness of VR-CAP versus R-CHOP instead of the European Union subgroup. The treatment effectiveness in the LYM-3002 trial seems lower for the European Union population and thus using this subgroup will result in more conservative estimates.

In general, the ERG has several comments on the way the PFS, OS and TFI were estimated. First, it is uncertain whether the best fitted model was selected to model the PFS for the VR-CAP arm. The log-logistic distribution was selected whereas the exponential distribution showed to have the best statistical fit. Another concern regarding the estimation of PFS is the use of a stratified model whereas an unstratified model seemed to be more appropriate. Second, regarding the modelling of survival, the ERG questions the use of different survival curves based on progression status and the assumption that survival for patients without progression differs between treatment arms. Third, the ERG expects that TFI is overestimated

by the company and possibly this overestimation is larger in VR-CAP patients since these estimates are based on all patients irrespective of treatment response.

The ERG finds the exclusion of AE during second-line treatment conservative as patients spent least time in the second-line when treated with VR-CAP and hence a greater negative impact of AE associated with second-line treatment would be expected for comparators. The ERG identified a model error in the calculation of AE cycle probabilities. However, these issues had only little impact on the ICERs.

Utilities for second-line treatment were based on assumptions and literature. The ERG agrees with the assumption that the utility for 'PFS from second-line treatment' is similar to the 'PFS from first-line treatments' however does not agree with the value used for 'progressed disease from second-line treatment'. This utility is based on a study about aggressive NHL and is clearly uncertain given the small patient population on which this estimate is based. The ERG also noted that some utility values used for the secondary scenario analyses are not correctly entered into the model of the company.

The ERG agreed with using the mean second-line duration for drug acquisition and administration as this was seen as conservative compared to using the maximum second-line duration. In contrast, the ERG did not agree with the dose reduction applied to the drug costs for VR-CAP and R-CHOP since it is questionable whether the dose reduction observed in the LYM-3002 trial are representative for UK clinical practice. In addition, some errors in the model were found regarding concomitant medication costs and costs for pegfilgrastim for R-CHOP.

All in all, some of the issues mentioned above were deemed of such importance that a new ERG base case was defined. In the company base case the ICERs compared with VR-CAP ranged between £13,725 (versus R-bendamustine) and £20,264 (versus R-CHOP). Once the ERG had implemented all changes, the ICERs compared with VR-CAP ranged between £13,455 (versus R-FC) and £34,039 (versus R-CHOP). Whilst the ICERs compared to R-FC and R-bendamustine were hardly influenced by the ERG changes, the ICER of VR-CAP versus R-CHOP increased by about 75%. In addition, the PSA showed that the probability that VR-CAP is cost-effective is much smaller in the ERG base-case compared to the company's base-case (11% versus 49% and 39% versus 89% for a threshold of £20,000 and £30,000, respectively). Similar to the company's base case, the probabilities that R-FC or R-bendamustine are cost-effective at the usual NICE thresholds are negligible.

In the exploratory ERG analyses the most extreme ICER was observed when the OS of VR-CAP was assumed to be similar to the OS of R-CHOP: £328,757. The next highest ICER was found if R-maintenance is included in the model, yielding an ICER of £43,779.

7. Impact on the ICER of additional clinical and economic analyses undertaken by the ERG

In section 6.3 the ERG base case was presented, which was based on various changes compared to the company base case. Tables 7.1, 7.2 and 7.3 show how each individual change impacts the ICER plus the combined effect of all changes simultaneously. From this we see that the largest impact on the change in ICER comes from changing the distribution for PFS in the VR-CAP arm to the exponential distribution, whilst keeping the distribution for R-CHOP PFS a log-logistic distribution.

Table 7.1: Revised base case cost-effectiveness analysis, incorporating corrections and amendments identified by the ERG: VR-CAP versus R-CHOP

	VR-CAP		R-CHOP		Incremental		ICER
	QALYs	Costs	QALYs	Costs	QALYs	Costs	Cost per QALY gained
Company base-case	**4.14**	**£45,724**	**3.34**	**£29,491**	**0.80**	**£16,234**	**£20,264**
Correction unit prices	4.12	£46,210	3.33	£29,979	0.79	£16,231	**£20,460**
Correction error adverse events	4.12	£45,808	3.32	£29,502	0.80	£16,306	**£20,322**
Correction costs concomitant medication	4.13	£46,137	3.33	£29,714	0.80	£16,423	**£20,507**
Age, weight and unit prices were not included as stochastic parameters	4.13	£45,684	3.34	£29,461	0.79	£16,223	**£20,514**
Proportion of patients receiving treatment included as stochastic parameter	4.17	£45,489	3.37	£29,199	0.80	£16,290	**£20,459**
Include half-cycle correction	4.13	£43,795	3.32	£28,289	0.81	£15,507	**£19,224**
PFS adjusted according to HR for European Union	4.08	£45,860	3.34	£29,398	0.74	£16,462	**£22,144**
Start second line treatment at time of progression	4.08	£46,267	3.31	£29,799	0.77	£16,467	**£21,354**
Utility progression 2nd line treatment based upon disutility Doorduijn	4.54	£45,624	3.87	£29,366	0.67	£16,257	**£24,190**
Per-protocol dosage	4.14	£48,838	3.35	£29,483	0.79	£19,355	**£24,410**
Primary assessment of progression	4.12	£45,521	3.39	£29,485	0.73	£16,035	**£21,961**
Survival is not distinguished between patients with and without progression	3.95	£42,843	3.24	£27,228	0.71	£15,614	**£21,987**
Exclude all-cause mortality	4.60	£45,871	3.52	£29,801	1.08	£16,070	**£14,818**
Exponential distribution for PFS in the VR-CAP arm	3.90	£47,025	3.37	£29,436	0.53	£17,589	**£33,087**
Indirect treatment comparison for efficacy R-FC and R-bendamustine	4.13	£45,601	3.34	£29,470	0.80	£16,130	**£20,282**
Exclude end-of life costs	4.11	£41,166	3.32	£24,672	0.79	£16,493	**£20,794**
ERG base case	**4.57**	**£44,140**	**3.93**	**£22,186**	**0.64**	**£21,955**	**£34,039**

Table 7.2: Revised base case cost-effectiveness analysis, incorporating corrections and amendments identified by the ERG: VR-CAP versus R-FC

	VR-CAP		R-FC		Incremental		ICER
	QALYs	Costs	QALYs	Costs	QALYs	Costs	Cost per QALY gained
Company base-case	**4.14**	**£45,724**	**3.36**	**£31,222**	**0.79**	**£14,503**	**£18,430**
Correction unit prices	4.12	£46,210	3.34	£31,863	0.78	£14,348	**£18,424**
Correction error adverse events	4.12	£45,808	3.34	£31,236	0.79	£14,572	**£18,492**
Correction costs concomitant medication	4.13	£46,137	3.34	£31,586	0.79	£14,551	**£18,506**
Age, weight and unit prices were not included as stochastic parameters	4.13	£45,684	3.35	£31,208	0.78	£14,475	**£18,643**
Proportion of patients receiving treatment is included as stochastic parameter	4.17	£45,489	3.38	£30,911	0.78	£14,578	**£18,641**
Include half-cycle correction	4.13	£43,795	3.34	£29,855	0.79	£13,940	**£17,596**
PFS adjusted according to HR for European Union	4.08	£45,860	3.35	£31,148	0.73	£14,713	**£20,178**
Start second line treatment at time of progression	4.08	£46,267	3.31	£31,670	0.77	£14,567	**£18,949**
Utility progression 2nd line treatment based upon disutility Doorduijn	4.54	£45,624	3.88	£31,095	0.67	£14,528	**£21,798**
Per-protocol dosage	4.14	£48,838	3.36	£31,218	0.78	£17,620	**£22,632**
Primary assessment of progression	4.12	£45,521	3.40	£31,224	0.71	£14,297	**£20,025**
Survival is not distinguished between patients with and without progression	3.95	£42,843	3.25	£29,038	0.70	£13,804	**£19,742**
Exclude all-cause mortality	4.60	£45,871	3.53	£31,543	1.07	£14,329	**£13,391**
Exponential distribution for PFS in the VR-CAP arm	3.90	£47,025	3.38	£31,177	0.52	£15,848	**£30,622**
Indirect treatment comparison for efficacy R-FC and R-bendamustine	4.13	£45,601	2.58	£29,440	1.56	£16,160	**£10,361**
Exclude end-of life costs	4.11	£41,166	3.33	£26,408	0.78	£14,758	**£18,960**
ERG base case	**4.57**	**£44,140**	**2.96**	**£22,370**	**1.62**	**£21,770**	**£13,455**

Table 7.3: Revised base case cost-effectiveness analysis, incorporating corrections and amendments identified by the ERG: VR-CAP versus R-bendamustine

	VR-CAP		R-bendamustine		Incremental		ICER
	QALYs	Costs	QALYs	Costs	QALYs	Costs	Cost per QALY gained
Company base-case	**4.14**	**£45,724**	**3.35**	**£34.883**	**0.79**	**£10,842**	**£13,725**
Correction unit prices	4.12	£46,210	3.34	£35,391	0.78	£10,819	**£13,838**
Correction error adverse events	4.12	£45,808	3.33	£34,867	0.79	£10,941	**£13,831**
Correction costs concomitant medication	4.13	£46,137	3.34	£35,180	0.79	£10,957	**£13,880**
Age, weight and unit prices were not included as stochastic parameters	4.13	£45,684	3.35	£34,847	0.78	£10,836	**£13,902**
Proportion of patients receiving treatment is included as stochastic parameter	4.17	£45,489	3.38	£34,603	0.79	£10,886	**£13,866**
Include half-cycle correction	4.13	£43,795	3.33	£33,157	0.80	£10,639	**£13,377**
PFS adjusted according to HR for European Union	4.08	£45,860	3.35	£34,806	0.73	£11,055	**£15,099**
Start second line treatment at time of progression	4.08	£46,267	3.31	£35,324	0.77	£10,943	**£14,222**
Utility progression 2nd line treatment based upon disutility Doorduijn	4.54	£45,624	3.88	£34,747	0.67	£10,877	**£16,288**
Per-protocol dosage	4.14	£48,838	3.36	£34,873	0.78	£13,965	**£17,867**
Primary assessment of progression	4.12	£45,521	3.40	£34,853	0.72	£10,668	**£14,872**
Survival is not distinguished between patients with and without progression	3.95	£42,843	3.25	£32,676	0.70	£10,166	**£14,490**
Exclude all-cause mortality	4.60	£45,871	3.53	£35,191	1.07	£10,680	**£ 9,953**
Exponential distribution for PFS in the VR-CAP arm	3.90	£47,025	3.38	£34,821	0.52	£12,204	**£23,446**
Indirect treatment comparison for efficacy R-FC and R-bendamustine	4.13	£45,601	3.64	£30,422	0.50	£15,179	**£30,398**
Exclude end-of life costs	4.11	£41,166	3.33	£30,063	0.78	£11,103	**£14,208**
ERG base case	**4.57**	**£44,140**	**3.43**	**£24,957**	**1.14**	**£19,184**	**£16,762**

8. Conclusions

8.1 Statement of principle findings

For the overall population, most efficacy outcomes favoured VR-CAP, but analyses were only significant for progression free survival and time to new treatment. For the Western Europe and North America only four analyses were reported, these all favoured VR-CAP but were not significant. Other outcome analyses were unavailable; more comprehensive results will be available with the final analysis in 2017.

The base case incremental cost-effectiveness ratio (ICER) for VR-CAP compared to R-CHOP was £20,362 per quality adjusted life year gained. ICERs versus other comparators were £18,509 versus R-FC and £13,797 versus R-bendamustine. These results were generally robust under the sensitivity and scenario analyses conducted by the Company, with no scenarios bringing the ICER of VR-CAP versus R-CHOP above £30,000. The ICER is most sensitive to variations within the fit of the parametric models for PFS and OS. The utility of patients who have progressed from second-line treatment is the parameter with the second greatest ICER sensitivity, due to patients spending a relatively long time in this health state.

A new base case was identified by the ERG, increasing the ICER of VR-CAP versus R-CHOP by £14,000 to £34,039 compared to the CS base case, which represents a 75% increase. The ICER of VR-CAP versus R-FC changed from £18,430 to £13,455 and the ICER of VR-CAP versus R-bendamustine increased from £13,725 to £16,762. The PSA showed that the probability that VR-CAP is cost-effective at a threshold of £20,000 and £30,000 is much smaller in the ERG base-case compared to the company's base-case (11% versus 49% and 39% versus 89% for a threshold of £20,000 and £30,000, respectively). Similar to the company's base case, the probabilities that R-FC or R-bendamustine are cost-effective at the usual NICE thresholds are negligible.

This large difference between the company base case and the ERG base case was caused mainly by changing the distribution for PFS in the VR-CAP arm to the exponential distribution, whilst keeping the distribution for R-CHOP PFS a log-logistic distribution.

In the exploratory ERG analyses that were done using the ERG base case as starting point, the most extreme ICER was observed when the overall survival of VR-CAP was assumed to be similar to the overall survival of R-CHOP: £328,757. The next highest ICER was found when R-maintenance is included in the model, yielding an ICER of £43,779 for VR-CAP versus R-CHOP.

8.2 Strengths and limitations of the assessment

Overall, the CS is well presented and in line with the final scope.

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 submission and response to clarification provided sufficient details for the ERG to appraise the searches. Additional searches were carried out for conference abstracts. However, it is not certain whether all relevant studies have been correctly identified and assessed as in initial phase only a single reviewer screened the retrieved references for potentially relevant studies. The ERG has identified three references,

linked to the LYM-3002 trial, which were incorrectly excluded but would not have changed the main conclusions of the CS.[32-34]

The three included RCTs (LYM-3002 for the comparison of VR-CAP with R-CHOP; European MCL Elderly trial and StiL NHL for indirect comparisons) were judged to be of high risk of bias (see Section 5.1.4).[9, 25, 29] Therefore, the results of these studies should be interpreted with caution. Furthermore, the three studies included relatively few participants with MCL (n=1067) which could limit the reliability of the findings. One of the trials (LYM-3002) is ongoing. The data used in the CS came from an interim analysis which was conducted in December 2013. It is likely that data on more participants will be available for the final analysis in 2017.

For the only RCT on VR-CAP, LYM-3002, no UK participants were included and approximately one third of the participants were from Europe or North America (Canada, USA).[9] Given the different prevalence depending on the geographic region and potential differences in clinical standards (e.g. concomitant care), the question of generalisability arises.[35, 36] When focussing on participants from Europe and North America, results for all endpoints will be underpowered.

The ERG was concerned about specific adverse events searches without the restriction of a study design filter were not conducted; this is not in line with current best practice. Searches of the Cochrane Library for sections 5.4 and 5.5 were overly restrictive and may have impaired recall. The dates of clinical effectiveness searches were reported inconsistently in Appendices 3 and 15 of the company submission.

According to the guidelines quoted by the CS[6], options for chemotherapy in MCL patients whom HSCT is unsuitable include R-CHOP, R-CVP, R-FC, R-bendamustine and R-chlorambucil. This indicates a greater number of alternative treatments than those discussed in the CS section of 'clinical pathways of care'. It is noted from the advisory board summary report[7] that R-CHOP is the preferred first-line treatment in the UK because it has the best evidence base to support its clinical benefit. Treatment options that were outside the scope for this project, namely R-CVP and R-chlorambucil, could potentially have allowed a wider network with more indirect comparisons.

Most uncertainty in the health economic model was related to the estimation of overall survival. The ERG questions the validity of the approach to separately model OS for patients with and without progression and the assumption that initial treatment only impacts OS for patients without progression.

8.3 Suggested research priorities

Currently, the evidence base for treatments of previously untreated patients with mantle cell lymphoma is limited. Best available evidence on VR-CAP comes from an interim analysis of a single RCT which does not include any patients from the UK.

Recent reviews recognise the heterogeneity of the disease (a subset of patients have indolent clinical course) and suggest that therapies should be tailored to the genetic/ cellular characteristics as well as the clinical characteristics of the disease.[8] Future research should

assess the clinical value of new and existing drugs for patients with MCL, depending on genetic and cellular characteristics.

Future research supporting NICE decisions should ideally include participants from the UK and/or a high number of patients from comparable geographical regions or genetic background. By including further relevant comparators, it would be possible to generate further evidence (from indirect comparisons and mixed treatment comparisons) to inform decision-making.

9. References

[1] National Institute for Health and Care Excellence. *Proposed Health Technology Appraisal: Bortezomib for previously untreated mantle cell lymphoma (Final scope). Appendix B [Internet]*. London: NICE, 2015 [accessed 16.7.15]. 3p. Available from: https://www.nice.org.uk/guidance/GID-TAG495/documents/lymphoma-mantle-cell-untreated-bortezomib-id724-final-scope2

[2] Janssen-Cilag International NV. Single technology appraisal. Lymphoma (mantle cell, untreated) – bortezomib [ID724] - company evidence submission to National Institute of Health and Care Excellence. Janssen-Cilag International NV., 26 June 2015: 206.

[3] Cancer Research UK. Non-Hodgkin lymphoma incidence statistics: Non-Hodgkin Lymphoma (NHL) incidence by sex and UK region [Internet]. London: Cancer Research UK, 2014 [accessed 21.7.15]. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/non-hodgkin-lymphoma/incidence#heading-Zero

[4] National Institute for Health Research Horizon Scanning Centre. *Bortezomib (Velcade) for newly diagnosed mantle cell lymphoma – first line. NIHR HSC ID: 1942 [Internet]*. Birmingham: NIHR Horizon Scanning Centre, University of Birmingham, 2012 [accessed 16.7.15]. 7p. Available from: http://www.hsc.nihr.ac.uk/topics/bortezomib-velcade-for-newly-diagnosed-mantle-cell/

[5] Office for National Statistics (ONS). Mortality Statistics: Deaths Registered in 2012 (Series DR) Tables 1–4 and Tables 6–14 (Excel sheet) [Internet]. London: ONS, 2013 [accessed 5.8.15]. Available from: http://www.ons.gov.uk/ons/rel/vsob1/mortality-statistics--deaths-registered-in-england-and-wales--series-dr-/2012/dr-tables-2012.xls

[6] McKay P, Leach M, Jackson R, Cook G, Rule S. Guidelines for the investigation and management of mantle cell lymphoma. *Br J Haematol* 2012;159(4):405-26.

[7] Janssen-Cilag Ltd. *Velcade MCL HTA Advisory Board Summary*, 2014

[8] Jares P, Colomer D, Campo E. Genetic and molecular pathogenesis of mantle cell lymphoma: perspectives for new targeted therapeutics. *Nat Rev Cancer* 2007;7(10):750-62.

[9] Robak T, Huang H, Jin J, Zhu J, Liu T, Samoilova O, et al. Bortezomib-based therapy for newly diagnosed mantle-cell lymphoma. *N Engl J Med* 2015;372(10):944-53.

[10] 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

[11] National Institute for Health and Care Excellence. *Single Technology Appraisal: company evidence submission template [Internet]*. London: NICE, 2015 [accessed 2.7.15].

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

[12] National Institute for Health and Care Excellence. *Single Technology Appraisal: user guide for company evidence submission template [Internet]*. London: NICE, 2015 [accessed 2.7.15]. 60p. Available from: https://www.nice.org.uk/article/pmg24/resources/non-guidance-single-technology-appraisal-user-guide-for-company-evidence-submission-template-pdf

[13] National Institute for Health and Care Excellence. *Guide to the methods of technology appraisal 2013 [Internet]*. London: NICE, 2013 [accessed 2.7.15]. 93p. Available from: http://publications.nice.org.uk/guide-to-the-methods-of-technology-appraisal-2013-pmg9/introduction

[14] Scottish Intercollegiate Guidelines Network. Search filters [Internet]. Edinburgh: SIGN, 2014 [accessed 20.8.15]. Available from: http://www.sign.ac.uk/methodology/filters.html

[15] Janssen-Cilag International NV. Single technology appraisal. Lymphoma (mantle cell, untreated) – bortezomib [ID724] - company evidence submission: appendices. Janssen-Cilag International NV., June 2015: 264.

[16] Janssen-Cilag International NV. Lymphoma (mantle cell, untreated) – bortezomib [ID724]. Response to request for clarification from the ERG. Janssen-Cilag International NV.,, 2015: 32.

[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] Higgins JPT, Green S, eds. *Cochrane handbook for systematic reviews of interventions [Internet]*. Version 5.1.0 [updated March 2011]: The Cochrane Collaboration, 2011 [accessed 23.3.11]. Available from: http://www.cochrane-handbook.org/

[19] Drach J, Huang H, Samoilova O, Belch A, Farber CM, Bosly A, et al. Efficacy and safety of frontline bortezomib, rituximab, cyclophosphamide, doxorubicin, and prednisone (VR-CAP) vs R-CHOP in a subset of newly diagnosed mantle cell lymphoma (MCL) patients (pts) medically eligible for transplantation in the randomized Phase 3 LYM-3002 study (NCT00722137). In: 56th American Society of Hematology (ASH) Annual Meeting and Exposition; San Francisco, CA, USA, 6-9 December 2014: Abstract 3064.

[20] Robak T, Huang H, Jin J, Zhu J, Liu T, Samoilova O, et al. Bortezomib (Btz) dose intensity is the strongest predictor for overall survival (OS) in mantle cell lymphoma (MCL) patients (pts) not considered for transplantation, receiving frontline btz plus rituximab, cyclophosphamide, doxorubicin, and prednisone (VR-CAP) therapy in the phase 3 LYM-

3002 study. In: 56th American Society of Hematology (ASH) Annual Meeting and Exposition; San Francisco, CA, USA, 6-9 December 2014: Abstract 4412.

[21] Janssen Research & Development. *A randomized, open-label, multicenter Phase III study of the combination of rituximab, cyclophosphamide, doxorubicin, VELCADE, and prednisone (VcR-CAP) or rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP) in patients with newly diagnosed mantle cell lymphoma who are not eligible for a bone marrow transplant. (Clinical Study Report; EDMS-ERI-25850170:2.0) [Data on file provided with the company's submission]*, 27 May 2014. 1416p.

[22] Eve HE, Linch D, Qian W, Ross M, Seymour JF, Smith P, et al. Toxicity of fludarabine and cyclophosphamide with or without rituximab as initial therapy for patients with previously untreated mantle cell lymphoma: results of a randomised phase II study. *Leuk Lymphoma* 2009;50(2):211-5.

[23] Rule S, Burton C, Waleski J, Jack A, Seymour J. A randomized phase II study of fludarabine/cyclophosphamide +/- rituximab in patients with untreated mantle cell lymphoma. Abstract 198. *Ann Oncol* 2005;16(Suppl 5):197.

[24] Flinn IW, van der Jagt R, Kahl BS, Wood P, Hawkins TE, Macdonald D, et al. Randomized trial of bendamustine-rituximab or R-CHOP/R-CVP in first-line treatment of indolent NHL or MCL: the BRIGHT study. *Blood* 2014;123(19):2944-52.

[25] Kluin-Nelemans HC, Hoster E, Hermine O, Walewski J, Trneny M, Geisler CH, et al. Treatment of older patients with mantle-cell lymphoma. *N Engl J Med* 2012;367(6):520-31.

[26] Lenz G, Dreyling M, Hoster E, Wormann B, Duhrsen U, Metzner B, et al. Immunochemotherapy with rituximab and cyclophosphamide, doxorubicin, vincristine, and prednisone significantly improves response and time to treatment failure, but not long-term outcome in patients with previously untreated mantle cell lymphoma: results of a prospective randomized trial of the German Low Grade Lymphoma Study Group (GLSG). *J Clin Oncol* 2005;23(9):1984-92.

[27] Hoster E, Unterhalt M, Wormann B, Duhrsen U, Metzner B, Eimermacher H, et al. The addition of rituximab to first-line chemotherapy (R-CHOP) results in superior response rates, time to treatment failure and response duration in patients with advanced stage mantle cell lymphoma: long term results of a randomized GLSG trial [Poster Abstract 3049]. *Blood* 2008 [accessed 25.8.15];112:1048.

[28] Rule S, Smith P, Johnson PW, Bolam S, Follows GA, Gambell J, et al. The addition of rituximab to fludarabine and cyclophosphamide (FC) improves overall survival in newly diagnosed mantle cell lymphoma (MCL): Results of the randomised UK national cancer research institute (NCRI) trial. In: 53rd American Society of Hematology (ASH) Annual Meeting and Exposition; San Diego, CA, USA, 10-13 December 2011: Abstract 440.

[29] Rummel MJ, Niederle N, Maschmeyer G, Banat GA, von Grunhagen U, Losem C, et al. Bendamustine plus rituximab versus CHOP plus rituximab as first-line treatment for patients with indolent and mantle-cell lymphomas: an open-label, multicentre, randomised, phase 3 non-inferiority trial. *Lancet* 2013;381(9873):1203-10.

[30] Rummel M. *MCL subgroup analyses of the randomised Phase III study of the Study Group Indolent Lymphomas (StiL) [PowerPoint Presentation]*, 2014. 13p.

[31] Rummel MJ, Maschmeyer G, Ganser A, Heider A, von Grunhagen U, Losem C, et al. Bendamustine Plus Rituximab (B-R) Versus CHOP Plus Rituximab (CHOP-R) As First-Line Treatment in Patients with Indolent and Mantle Cell Lymphomas (MCL) – 7 Year Updated Results from the StiL NHL1 Study In: 56th American Society of Hematology (ASH) Annual Meeting and Exposition; San Francisco, CA, USA, 6-9 December 2014: Abstract 4407.

[32] Cavalli F, Rooney B, Pei L, van de Velde H, Robak T, on behalf of the LYM-3002 Investigators. Randomized phase 3 study of rituximab, cyclophosphamide, doxorubicin, and prednisone plus vincristine (R-CHOP) or bortezomib (VR-CAP) in newly diagnosed mantle cell lymphoma (MCL) patients (pts) ineligible for bone marrow transplantation (BMT). Abstract presented at ASCO Annual Meeting 2014 [Internet]. *J Clin Oncol* 2014 [accessed 26.8.15];32(5 (Suppl)):Abstract 8500. Available from: http://meetinglibrary.asco.org/content/129206-144

[33] Ogura M, Pei L, Rooney B, van de Velde H, Cavalli F. 3061. Efficacy and safety of bortezomib plus rituximab, cyclophosphamide, doxorubicin, and prednisone (VR-CAP) vs R-CHOP in japanese patients (pts) with newly diagnosed mantle cell lymphoma (MCL) not considered for transplantation enrolled in the international, randomized, phase 3 LYM-3002 study (NCT00722137). Abstract presented at 56th American Society of Hematology (ASH) Annual Meeting and Exposition; 6-9 December 2014; San Francisco: US [Internet]. 2014 [accessed 26.8.15]. Available from: https://ash.confex.com/ash/2014/webprogram/Paper69583.html

[34] Robak T, Huang H, Jin J, Zhu J, Liu T, Samoilova O, et al. Phase 3 study of frontline rituximab, cyclophosphamide, doxorubicin, and prednisone plus vincristine (R-CHOP) or bortezomib (VR-CAP) in transplantation-unsuitable mantle cell lymphoma (MCL) patients. *Clin Adv Hematol Oncol* 2014;12(8 (Suppl 16)):9-11.

[35] Anderson JR, Armitage JO, Weisenburger DD. Epidemiology of the non-Hodgkin's lymphomas: distributions of the major subtypes differ by geographic locations. Non-Hodgkin's Lymphoma Classification Project. *Ann Oncol* 1998;9(7):717-20.

[36] Chihara D, Ito H, Matsuda T, Shibata A, Katsumi A, Nakamura S, et al. Differences in incidence and trends of haematological malignancies in Japan and the United States. *Br J Haematol* 2014;164(4):536-45.

[37] Akiyama N, Tsuruta H, Sasaki H, Sakamoto H, Hamaguchi M, Ohmura Y, et al. Messenger RNA levels of five genes located at chromosome 11q13 in B-cell tumors with chromosome translocation t(11;14)(q13;q32). *Cancer Res* 1994;54(2):377-9.

[38] Song F, Loke YK, Walsh T, Glenny A, Eastwood AJ, Altman DG. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009;338:1147.

[39] Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997;50(6):683-91.

[40] Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess* 2004;8(36):iii-iv, ix-xi, 1-158.

[41] Doorduijn J, Buijt I, Holt B, Steijaert M, Uyl-de Groot C, Sonneveld P. Self-reported quality of life in elderly patients with aggressive non-Hodgkin's lymphoma treated with CHOP chemotherapy. *Eur J Haematol* 2005;75(2):116-23.

[42] European Medicines Agency (EMA). *Velcade - bortezomib: EPAR summary for the public [Internet]*. London: European Medicines Agency (EMA), 2015 [accessed 16.7.15]. 4p. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Summary_for_the_public/human/000539/WC500048136.pdf

[43] European Medicines Agency. *Annex 1. Summary of product characteristics (SPC): VELCADE 1 mg powder for solution for injection [Internet]*. London: EMA, 2015 [accessed 1.9.15]. 101p. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000539/WC500048471.pdf

[44] Janssen-Cilag International NV. *Summary of product characteristics (SPC): Velcade 3.5 mg powder for solution for injection [Internet]*, 2015 [accessed 16.7.15] Available from: http://www.medicines.org.uk/EMC/medicine/17109/SPC/Velcade+3.5mg+powder+for+solution+for+injection/

[45] Amin A, Taylor R, Alonso A, Schwenke C, Gaudig M, Gaugris S, et al. Progression-free survival as a surrogate endpoint for overall survival in relapsed-refractory mantle cell lymphoma. In: 19th European Hematology Association Congress; Milan, Italy, 12-15 June 2014: Abstract P456.

[46] Abrahamsson A, Albertsson-Lindblad A, Brown PN, Baumgartner-Wennerholm S, Pedersen LM, D'Amore F, et al. Real world data on primary treatment for mantle cell lymphoma: a Nordic Lymphoma Group observational study. *Blood* 2014;124(8):1288-95.

[47] National Institute for Health and Care Excellence. *Rituximab for the first-line maintenance treatment of follicular non-Hodgkin's lymphoma. NICE technology appraisal guidance 226 [Internet]*. London: NICE, 2011 [accessed 28.8.15] Available from: http://www.nice.org.uk/guidance/TA226

[48] Latimer N. *National Institute for Health and Clinical Excellence (NICE) DSU technical support document 14: survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data [Internet]*. Sheffield: Decision Support Unit, ScHARR, University of Sheffield., 2011 (last updated March 2013) [accessed 16.7.15]. 52p. Available from: http://www.nicedsu.org.uk/NICE%20DSU%20TSD%20Survival%20analysis.updated%20March%202013.v2.pdf

[49] Lee L, Wang L, Crump M. Identification of potential surrogate end points in randomized clinical trials of aggressive and indolent non-Hodgkin's lymphoma: correlation of complete response, time-to-event and overall survival end points. *Ann Oncol* 2011;22(6):1392-403.

[50] Uyl-de Groot CA, Hagenbeek A, Verdonck LF, Lowenberg B, Rutten FF. Cost-effectiveness of ABMT in comparison with CHOP chemotherapy in patients with intermediate- and high-grade malignant non-Hodgkin's lymphoma (NHL). *Bone Marrow Transplant* 1995;16(3):463-70.

[51] van Agthoven M, Vellenga E, Fibbe WE, Kingma T, Uyl-de Groot CA. Cost analysis and quality of life assessment comparing patients undergoing autologous peripheral blood stem cell transplantation or autologous bone marrow transplantation for refractory or relapsed non-Hodgkin's lymphoma or Hodgkin's disease. a prospective randomised trial. *Eur J Cancer* 2001;37(14):1781-9.

[52] Witzens-Harig M, Reiz M, Heiss C, Benner A, Hensel M, Neben K, et al. Quality of life during maintenance therapy with the anti-CD20 antibody rituximab in patients with B cell non-Hodgkin's lymphoma: results of a prospective randomized controlled trial. *Ann Hematol* 2009;88(1):51-7.

[53] Crott R, Versteegh M, Uyl-de-Groot C. An assessment of the external validity of mapping QLQ-C30 to EQ-5D preferences. *Qual Life Res* 2013;22(5):1045-54.

[54] Soini EJ, Martikainen JA, Vihervaara V, Mustonen K, Nousiainen T. Economic evaluation of sequential treatments for follicular non-hodgkin lymphoma. *Clin Ther* 2012;34(4):915-925 e2.

[55] Hayslip JW, Simpson KN. Cost-effectiveness of extended adjuvant rituximab for US patients aged 65-70 years with follicular lymphoma in second remission. *Clin Lymphoma Myeloma* 2008;8(3):166-70.

[56] Deconinck E, Miadi-Fargier H, Pen CL, Brice P. Cost effectiveness of rituximab maintenance therapy in follicular lymphoma: long-term economic evaluation. *Pharmacoeconomics* 2010;28(1):35-46.

[57] Addicot R, Dewar S. *Improving choice at the end of life: a descriptive analysis of the impact and costs of the Marie Curie Delivering Choice programme in Lincolnshire [Internet]*. London: King's Fund, 2008 [accessed 28.8.15] Available from: http://www.kingsfund.org.uk/sites/files/kf/improving-choice-end-of-life-descriptive-analysis-impact-costs-marie-curie-choice-programme-lincolnshire-rachael-addicot-steve-dewar-april-2008.pdf

[58] Dept of General Practice, University of Glasgow. *Critical appraisal checklist for economic evaluations [Internet]*. Glasgow: University of Glasgow, 1997 [accessed 28.8.15] Available from: http://www.gla.ac.uk/media/media_64048_en.pdf

Appendix 1: ERG search strategies

Clinical effectiveness

The dates of clinical effectiveness searches were reported inconsistently in Appendices 3 and 15 of the company submission.[15] These discrepancies were not clarified in the clarification response.[16]

Cost-effectiveness

No database provider was reported for EconLit.

Measurement and value of health effects

The Cochrane Library strategy in Appendix 13 (pages 144-146) incorporated incorrect truncation syntax for the Wiley interface.[15] Ovid post-truncation qualification had been applied which resulted in zero results, e.g. "lymphoplasmacytic lymphoma*1:ti,ab". The Wiley interface does not support specification of truncation in this way, therefore each line that included this command has not worked correctly. The ERG corrected and re-ran the NHS EED search, and screened the missed references.

Cost and healthcare resource use identification, measurement and valuation

Appendix 14 of the CS incorrectly reported that these searches included terms for NHL.[15] This was confirmed by the company in their clarification response.[16]

Additional ERG search of NHS EED for 5.4 Measurement and valuation of health effects

NHS Economic Evaluation Database (NHS EED) (Wiley): Issue 2/April 2015: up to 2015/08/18

Searched 18.8.15

#1	MeSH descriptor: [Lymphoma, Non-Hodgkin] explode all trees	1254
#2	(non hodgkin* or nonhodgkin* or nhl or nhls):ti,ab	2039
#3	(follicular lymphoma* or fl):ti,ab	1435
#4	(small lymphocytic lymphoma* or sll):ti,ab	70
#5	MeSH descriptor: [Waldenstrom Macroglobulinemia] this term only	16
#6	(waldenstrom* near/1 (macroglobulin?emia* or lymphoma*)):ti,ab	4
#7	lymphoplasmacytic lymphoma*:ti,ab	20
#8	(marginal zone lymphoma* or mzl):ti,ab	42
#9	MeSH descriptor: [Lymphoma, Mantle-Cell] this term only	48
#10	(mantle cell or MCL):ti,ab	238
#11	((bcell* or b cell*) near/2 lymphoma*):ti,ab	369
#12	#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11	4116

NHS EED search retrieved 59 results.

Additional ERG search of NHS EED for 5.5 Cost and health care resource identification, measurement and valuation

NHS Economic Evaluation Database (NHS EED) (Wiley): Issue 2/April 2015: up to 2015/08/18

Searched 18.8.15

#1	MeSH descriptor: [Lymphoma, Mantle-Cell] this term only	48

#2	(mantle cell or MCL):ti,ab	238
#3	((bcell* or b cell*) near/2 lymphoma*):ti,ab	369
#4	#1 or #2 or #3	593

NHS EED search retrieved 5 records.

Appendix 2: Summary list of cost-effectiveness evaluations

Question(s)	***Response (Y, N or NS)***	***Comments***
Is there a clear statement of the decision problem?	Y	
Is the objective of the evaluation and model specified and consistent with the stated decision problem?	Y	
Is the primary decision-maker specified?	Y	
Is the perspective of the model stated clearly?	Y	NHS-PPS
Are the model inputs consistent with the stated perspective?	Partially	End-of-life costs include costs for 'third sector' (i.e. not-for-profit and non-governmental) healthcare organisations. Moreover, the LYM-3002 trial included no patients from the UK and only a few from the EU.
Has the scope of the model been stated and justified?	Y	
Are the outcomes of the model consistent with the perspective, scope and overall objective of the model?	Y	
Is the structure of the model consistent with a coherent theory of the health condition under evaluation?	Y	The model distinguishes patients with and without disease progression. Progression is considered to be an important factor regarding prognosis of patients with MCL Furthermore, the model distinguishes different treatment lines.
Are the sources of data used to develop the structure of the model specified?	N	A new model has been developed and it is only stated that progression is a surrogate of survival. No specific documentation for the choices made regarding model structure
Are the causal relationships described by the model structure justified appropriately?	Y	
Are the structural assumptions transparent and justified?	Y	
Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Y	
Is there a clear definition of the options under evaluation?	Y	
Have all feasible and practical options been evaluated?	Y	
Is there justification for the exclusion of feasible options?	NA	In principle, all options have been evaluated. However, some options are only evaluated in a

Question(s)	***Response (Y, N or NS)***	***Comments***
		scenario analysis as these are considered no real comparator for VR-CAP as these will probably administered if R-CHOP and VR-CAP are not feasible
Is the chosen model type appropriate given the decision problem and specified causal relationships within the model?	Y	
Is the time horizon of the model sufficient to reflect all important differences between options?	Y	
Are the time horizon of the model, the duration of treatment and the duration of treatment effect described and justified?	Y	Duration of treatment effect is assumed to be lifelong due to the method of estimation
Do the disease states (state transition model) or the pathways (decision tree model) reflect the underlying biological process of the disease in question and the impact of interventions?	Y	
Is the cycle length defined and justified in terms of the natural history of disease?	Y	
Are the data identification methods transparent and appropriate given the objectives of the model?	Y	
Where choices have been made between data sources, are these justified appropriately?	Y	
Has particular attention been paid to identifying data for the important parameters in the model?	Y	Literature search have been done regarding inputs of QoL and costs
Has the quality of the data been assessed appropriately?	Y	Quality assessments of the studies used for effectiveness have been assessed. However, no real assessment regarding the quality of costs and quality of life input
Where expert opinion has been used, are the methods described and justified?	N	Several choices have been made based upon expert opinion, but no detailed information is presented how the opinion was extracted.
Is the data modelling methodology based on justifiable statistical and epidemiological techniques?	Y	
Is the choice of baseline data	Y	

Question(s)	***Response (Y, N or NS)***	***Comments***
described and justified?		
Are transition probabilities calculated appropriately?	N	The distribution with the best fit to the data was not selected for the PFS of the VR-CAP arm. The probability to start second line treatment after progression from first line treatment is not correctly calculated: should be derived from patients with progression, not from any patients starting first line treatment.
Has a half-cycle correction been applied to both cost and outcome?	N	
If not, has this omission been justified?	N	
If relative treatment effects have been derived from trial data, have they been synthesised using appropriate techniques?	N	Only treatment effects from individual studies were used in the assessment of the impact of maintenance treatment.
Have the methods and assumptions used to extrapolate short-term results to final outcomes been documented and justified?	Y	
Have alternative extrapolation assumptions been explored through sensitivity analysis?	Y	Different statistical distributions were evaluated for the following endpoints: OS, PrePS, PPS, TFI
Have assumptions regarding the continuing effect of treatment once treatment is complete been documented and justified?	NA	
Have alternative assumptions regarding the continuing effect of treatment been explored through sensitivity analysis?	NA	
Are the costs incorporated into the model justified?	Y	
Has the source for all costs been described?	N	Not all unit prices could be reproduced (e.g. haematological visits)
Have discount rates been described and justified given the target decision-maker?	Y	
Are the utilities incorporated into the model appropriate?	N	Utilities were directly derived from the LYM-3002 trial for two health states. Assumptions and other sources were used for the other health states. It is questionable whether these are appropriate

Question(s)	***Response (Y, N or NS)***	***Comments***
Is the source for the utility weights referenced?	N	
Are the methods of derivation for the utility weights justified?	Y	
Have all data incorporated into the model been described and referenced in sufficient detail?	N	It was not completely clear how the utilities from literature were derived.
Has the use of mutually inconsistent data been justified (i.e. are assumptions and choices appropriate)?	Y	
Is the process of data incorporation transparent?	Y	
If data have been incorporated as distributions, has the choice of distribution for each parameter been described and justified?	N	
If data have been incorporated as distributions, is it clear that second order uncertainty is reflected?	N	Age and weight are also incorporated in the PSA (these parameters do not reflect second order uncertainty).
Have the four principal types of uncertainty been addressed?	N	No assessment of structural and methodological uncertainty
If not, has the omission of particular forms of uncertainty been justified?	N	
Have methodological uncertainties been addressed by running alternative versions of the model with different methodological assumptions?	N	
Is there evidence that structural uncertainties have been addressed via sensitivity analysis?	N	
Has heterogeneity been dealt with by running the model separately for different subgroups?	N	
Are the methods of assessment of parameter uncertainty appropriate?	Y	
If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly and justified?	N	Not always clear what they used as ranges for the sensitivity analysis
Is there evidence that the mathematical logic of the model has been tested thoroughly before use?	N	

Question(s)	***Response (Y, N or NS)***	***Comments***
Are any counterintuitive results from the model explained and justified?	NA	
If the model has been calibrated against independent data, have any differences been explained and justified?	NA	
Have the results of the model been compared with those of previous models and any differences in results explained?	NA	No other models/studies available

Appendix 3: Details of model changes implemented by the ERG

Issue	Sheet	Cell/name	Old formula	New formula	Comment
Correction unit costs cyclophophamide oral	Costs	c_cyclo_oral	£0.82	£1.39	
Correction unit costs administration by district nurse	Costs	c_districtnurse_unit	£65	£76	
Correction unit cost haematological visit	Costs	c_haematologist_unit	£150.06	£154.66	
Correction unit costs moxifloxacin	Costs	c_moxi	£1.90	£2.09	
Correction calculation adverse events	Adverse events	E74	0	=Parameters!P379	
Correction calculation adverse events	Adverse events	E75	0	=Parameters!P380	
Correction calculation concomitant medication	Resource Use	D106	=F70*c_filgras_week+G70*c_pegfil_week+H70*c_leno_week+F82*c_cipro_week+G82*c_levo_week+H82*c_moxi_week+I82*c_cefta_week+J82*c_ceftria_week+K82*c_meropenem_week+L82*c_coamox_week+M82*c_amox_week+N82*c_piptaz_week+O82*c_cotrim_week+P82*c_vanco_week+Q82*c_amika_week+F97*c_aciclo_week+G97*c_valaciclo_week	=IF(control_cost_conc="No";(F70*c_filgras_week+G70*c_pegfil_week+H70*c_leno_week)*(226/137)+(F82*c_cipro_week+G82*c_levo_week+H82*c_moxi_week+I82*c_cefta_week+J82*c_ceftria_week+K82*c_meropenem_week+L82*c_coamox_week+M82*c_amox_week+N82*c_piptaz_week+O82*c_cotrim_week+P82*c_vanco_week+Q82*c_amika_week)*(578/369)+(F97*c_aciclo_week+G97*c_valaciclo_week)*(184/159);0)	Copy until cell D111

Issue	Sheet	Cell/name	Old formula	New formula	Comment
Correction calculation concomitant medication	Resource Use	E106	=K70*c_filgras_week+L70*c_filgras_week+M70*c_leno_week+T82*c_cipro_week+U82*c_levo_week+V82*c_moxi_week++W82*c_cefta_week+X82*c_ceftria_week+Y82*c_meropenem_week+Z82*c_coamox_week+AA82*c_amox_week+AB82*c_piptaz_week+AC82*c_cotrim_week+AD82*c_vanco_week+AE82*c_amika_week+J97*c_aciclo_week+K97*c_valaciclo_week	=IF(control_cost_conc="No";(K70*c_filgras_week+L70*c_pegfil_week+M70*c_leno_week)*(171/128)+(T82*c_cipro_week+U82*c_levo_week+V82*c_moxi_week++W82*c_cefta_week+X82*c_ceftria_week+Y82*c_meropenem_week+Z82*c_coamox_week+AA82*c_amox_week+AB82*c_piptaz_week+AC82*c_cotrim_week+AD82*c_vanco_week+AE82*c_amika_week)*(400/259)+(J97*c_aciclo_week+K97*c_valaciclo_week)*(133/106);0)	Copy until cell E111
Include half-cycle correction	PF.VR-CAP	VRCAP_LY_Undisc	=SUM(AV12:AV1838)	=AV12*0.5+SUM(AV13:AV1838)	Copy to cells: VRCAP_LY_disc VRCAP_QALY_disc VRCAP_QALY_undisc VRCAP_Cost_undisc VRCAP_Costs_disc RCHOP_LY_undisc RCHOP_LY_disc RCHOP_QALY_disc RCHOP_QALY_undisc RCHOP_Cost_undisc RCHOP_Costs_disc RFC_LY_undisc

Issue	Sheet	Cell/name	Old formula	New formula	Comment
					RFC_LY_disc RFC_QALY_disc RFC_QALY_undisc RFC_Cost_undisc RFC_Costs_disc RBENDA_LY_undisc RBENDA_LY_disc RBENDA_QALY_disc RBENDA_QALY_undisc RBENDA_Cost_undisc RBENDA_Costs_disc
Age, Weight and unit prices were made fixed	Parameters	P17	=IF(C17="";"";IF(O17=1;D17;IF(O17=2;H17;IF(O17=3;I17;IF(O17=4;N17;""))))	=D17	Copy to P18, P254-P276, P280-P298
Adjust HR according to EU population	Progression & Survival	C318		=IF(control_EUpop="Yes";IF(control_assessment=Lists!$A$18;0.58/0.56;IF(control_assessment=Lists!$A$19;0.68/0.63;0.58/0.51));1)	Control_EUpop = new variable in control sheet with Yes/No choice
	Progression & Survival	D324	=$K324	=$K324^$C$318	Refers to cell C318 Copy down to end of the column
Start second line at time of progression	PF.VR-CAP	H13	=IF($C13>=Control_TimeHorizon_Days/7;"";(G12*Q13+H12*T13))	=IF($C13>=Control_TimeHorizon_Days/7;"";IF(secondline_rule="Progression";0;(G12*Q13+H12*T13)))	Copy down to the end of the column. Similar changes to the columns in sheet PF.R--CHOP, PF_R-FC, PF-R-

Issue	Sheet	Cell/name	Old formula	New formula	Comment
					benda
	PF.VR-CAP	I13	=IF($C13>=Control_TimeHorizon_Days/7;"";H12*U13+I12*X13)	=IF($C13>=Control_TimeHorizon_Days/7;"";IF(secondline_rule="Progression";G12*Q13;H12*U13)+I12*X13)	Copy down to the end of the column. Similar changes to the columns in sheet PF.R--CHOP, PF_R-FC, PF-R-benda
Change quality of life utilities for progression from second line	Lists	BR47	=IF(control_ut_2lpost=$A$44;BT54;IF(control_utilities=$A$22;IF(control_assessment=$A$18;IF(control_pop=$A$5;BT47;BU47);IF(control_assessment=$A$19;IF(control_pop=$A$5;BV47;BW47);IF(control_pop=$A$5;BX47;BY47)));IF(control_utilities=$A$23;BT54;IF(control_utilities=$A$24;BR44-(BV51-BV54);IF(control_utilities=$A$25;BR44-BW51-BW54;BR44-BX51-BX54)))))	=IF(control_ut_2lpost=$A$44;IF(control_disut="No";BR44+AVERAGE(BP65:BQ65);BR44+BP65);IF(control_utilities=$A$22;IF(control_assessment=$A$18;IF(control_pop=$A$5;BT47;BU47);IF(control_assessment=$A$19;IF(control_pop=$A$5;BV47;BW47);IF(control_pop=$A$5;BX47;BY47)));IF(control_utilities=$A$23;BT54;IF(control_utilities=$A$24;BR44-(BV51-BV54);IF(control_utilities=$A$25;BR44-BW51-BW54;BR44-BX51-BX54)))))	
Exclude end-of life costs	Costs	c_endoflife_unit	=5324*(D262/D256)	=0	
Include overall survival instead of dinstinction prePS and postPS	Progression & Survival	F323	=IF(control_combined_OS="Yes";$O323;$R323)	=IF(control_OS_Pstatus="Yes";IF(control_combined_OS="Yes";$O323;$R323);IF(control_combined_OS="YES";AVERAGE('Progression & Survival'!M323:N323);M323))	Add in control new variable: distinction control_OS_PStatus to identify whether overall survival of distinction pre and post progression survival. Copy down to the

Issue	Sheet	Cell/name	Old formula	New formula	Comment
					end fo the column
Include overall survival instead of dinstinction prePS and postPS	Progression & Survival	G323	=IF(control_combined_OS_preprog="No";$P323;AVERAGE(P323:Q323))	=IF(control_OS_Pstatus="Yes";IF(control_combined_OS_preprog="No";$P323;AVERAGE(P323:Q323));IF(control_combined_OS_preprog="No";M323;AVERAGE(M323:N323)))	Add in control new variable: distinction control_OS_PStatus to identify whether overall survival of distinction pre and post progression survival. Copy down to the end of the column+
Include overall survival instead of dinstinction prePS and postPS	Progression & Survival	H323	=IF(control_combined_OS="Yes";$O323;$S323)	=IF(control_OS_Pstatus="Yes";IF(control_combined_OS="Yes";$O323;$S323);IF(control_combined_OS="Yes";AVERAGE(M323:N323);N323))	Add in control new variable: distinction control_OS_PStatus to identify whether overall survival of distinction pre and post progression survival. Copy down to the end of the column
Include overall survival instead of dinstinction prePS and postPS	Progression & Survival	I323	=IF(control_combined_OS_preprog="No";$Q323;AVERAGE(P323:Q323))	=IF(control_OS_Pstatus="Yes";IF(control_combined_OS_preprog="No";$Q323;AVERAGE(P323:Q323));IF(control_combined_OS_preprog="No";N323;AVERAGE(M323:N323)))	Add in control new variable: distinction control_OS_PStatus to identify whether overall survival of distinction pre and post progression survival. (yes/no option) Copy down to the

Issue	Sheet	Cell/name	Old formula	New formula	Comment
					end of the column
Different functional form for PFS in the VR-CAP arm	Progerssion & Survival	K323	=IF(control_PFSindiv="Yes";IF(Control_Curve_PFS="Exponential";EXP(-1*EXP(-$C$257)*$C323);IF(Control_Curve_PFS="Weibull";EXP(-1*(($C323/EXP($C$262))^(1/$C$263)));IF(Control_Curve_PFS="Log-logistic";IF($C323=0;1;(1+EXP((LN($C323)-$C$267)/$C$268))^-1);IF(Control_Curve_PFS="Lognormal";IF($C323=0;1;(1-LOGNORMDIST($C323;$C$272;$C$273)));IF(Control_Curve_PFS="Gamma";IF(C323=0;1;IF($C$279>0;1-GAMMA.DIST($C$279^(-2)*(($C323*EXP(-$C$277))^(1/$C$278))^$C$279;$C$279^(-2);1;TRUE);GAMMA.DIST($C$279^(-2)*(($C323*EXP(-$C$277))^(1/$C$278))^$C$279;$C$279^(-2);1;TRUE)));IF(Control_Curve_PFS="Gompertz";EXP((EXP(-$C$283)/$C$284)*(1-EXP($C$284*($C323/(365.25/12)))));"Curve Selection Error"))))));IF(Control_Curve_PFS="Exponential";EXP(-1*EXP(-($C$57+$C$58))*$C323);IF(Control_C	=IF(control_PFSindiv="Yes";IF(Control_Curve_PFS_VRCAP="Exponential";EXP(-1*EXP(-$C$257)*$C323);IF(Control_Curve_PFS_VRCAP="Weibull";EXP(-1*(($C323/EXP($C$262))^(1/$C$263)));IF(Control_Curve_PFS_VRCAP="Log-logistic";IF($C323=0;1;(1+EXP((LN($C323)-$C$267)/$C$268))^-1);IF(Control_Curve_PFS_VRCAP="Lognormal";IF($C323=0;1;(1-LOGNORMDIST($C323;$C$272;$C$273)));IF(Control_Curve_PFS_VRCAP="Gamma";IF(C323=0;1;IF($C$279>0;1-GAMMA.DIST($C$279^(-2)*(($C323*EXP(-$C$277))^(1/$C$278))^$C$279;$C$279^(-2);1;TRUE);GAMMA.DIST($C$279^(-2)*(($C323*EXP(-$C$277))^(1/$C$278))^$C$279;$C$279^(-2);1;TRUE)));IF(Control_Curve_PFS_VRCAP="Gompertz";EXP((EXP(-$C$283)/$C$284)*(1-EXP($C$284*($C323/(365.25/12)))));"Curve Selection Error"))))));IF(Control_Curve_PFS_VRCAP="Exponential";EXP(-1*EXP(-($C$57+$C$58))*$C323);IF(Control_Curve_PFS_VRCAP="Weibull";EXP(-1*(($C323/EXP($C$63+$C$64))^(1/$C$65	Add in control new variable for functional form VR-CAP. Copy down to the end of the column

Issue	Sheet	Cell/name	Old formula	New formula	Comment
			urve_PFS="Weibull";EXP(-1*(($C323/EXP($C$63+$C$64))^(1/$C$65)));IF(Control_Curve_PFS="Log-logistic";IF($C323=0;1;(1+EXP((LN($C323)-($C$70+$C$71))/$C$72))^-1);IF(Control_Curve_PFS="Lognormal";IF($C323=0;1;(1-LOGNORMDIST($C323;($C$77+$C$78);$C$79)));IF(Control_Curve_PFS="Gamma";IF(C323=0;1;IF($C$87>0;1-GAMMA.DIST($C$87^(-2)*(($C323*EXP(-($C$84+$C$85)))^(1/$C$86))^$C$87;$C$87^(-2);1;TRUE);GAMMA.DIST($C$87^(-2)*(($C323*EXP(-($C$84+$C$85)))^(1/$C$86))^$C$87;$C$87^(-2);1;TRUE)));IF(Control_Curve_PFS="Gompertz";EXP((EXP(-($C$92+$C$93))/$C$94)*(1-EXP($C$94*($C323/(365.25/12)))));"Curve Selection Error")))))))	)));IF(Control_Curve_PFS_VRCAP="Log-logistic";IF($C323=0;1;(1+EXP((LN($C323)-($C$70+$C$71))/$C$72))^-1);IF(Control_Curve_PFS_VRCAP="Lognormal";IF($C323=0;1;(1-LOGNORMDIST($C323;($C$77+$C$78);$C$79)));IF(Control_Curve_PFS_VRCAP="Gamma";IF(C323=0;1;IF($C$87>0;1-GAMMA.DIST($C$87^(-2)*(($C323*EXP(-($C$84+$C$85)))^(1/$C$86))^$C$87;$C$87^(-2);1;TRUE);GAMMA.DIST($C$87^(-2)*(($C323*EXP(-($C$84+$C$85)))^(1/$C$86))^$C$87;$C$87^(-2);1;TRUE)));IF(Control_Curve_PFS_VRCAP="Gompertz";EXP((EXP(-($C$92+$C$93))/$C$94)*(1-EXP($C$94*($C323/(365.25/12)))));"Curve Selection Error")))))))	