The clinical and cost-effectiveness of Rituximab for the 1st line treatment of Chronic Lymphocytic Leukaemia: an evidence review of the submission from Roche

Report commissioned by: NHS R&D HTA Programme

On behalf of: NICE

Produced by: Peninsula Technology Assessment Group (PenTAG)

Authors: Caroline Main, Research Fellow, PenTAG
Martin Pitt, Senior Research Fellow, PenTAG
Tiffany Moxham, Information Scientist, PenTAG
Ken Stein, Professor of Public Health, Director of PenTAG

Correspondence to: [redacted]

Date completed: 30 January 2009

About The Peninsula Technology Assessment Group (PenTAG)

The Peninsula Technology Assessment Group is part of the Institute of Health Service Research at the Peninsula Medical School. PenTAG was established in 2000 and carries out independent Health Technology Assessments for the UK HTA Programme, systematic reviews and economic analyses for the NICE Centre for Public Health Excellence, and systematic reviews as part of the Cochrane Collaboration Heart Group, as well as for other local and national decision-makers. The group is multi-disciplinary and draws on individuals' backgrounds in public health, health services research, computing and decision analysis, systematic reviewing, statistics and health economics. The Peninsula Medical School is a school within the Universities of Plymouth and Exeter. The Institute of Health Research is made up of discrete but methodologically related research groups, among which Health Technology Assessment is a strong and recurring theme.

Source of funding

This report was commissioned by the NHS R&D HTA programme.

Competing interests of authors

None.

Acknowledgements

We particularly acknowledge the help of Claudius Rudin, Royal Devon and Exeter Hospital who provided clinical input during the project.

With many thanks to Joanne Perry (PenTAG) for administrative project support.

The views expressed in this report are those of the authors and not necessarily those of the NHS R&D HTA Programme. Any errors are the responsibility of the authors.

Contributions of authors

Caroline Main	Provided overall project management, critiqued the effectiveness evaluation provided by the manufacturers and contributed to the writing of the report
Martin Pitt	Critiqued the model provided by industry and contributed to the writing of the report

Tiffany Moxham	Undertook literature searches for the report, and commented on the searches provided by industry
Ken Stein	Contributed to the critique of the industry submission, report writing and editing

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AE	Adverse event
C	Cyclophosphamide
CHOP	Cyclophosphamide, doxorubicin, vincristine, prednisone
CI	Confidence interval
CLL	Chronic lymphocytic leukaemia
CLL-8	Trial number ML17102 led by the German Chronic Lymphocytic Leukaemia Study Group
CR	Complete response
DFS	Disease-free survival
ECOG	Eastern Cooperative Oncology Group
EFS	Event-free survival
EMEA	European Medicines Agency
ETR	End-of-treatment response
EORTC-QLQ-C30	European Organization for Research and Treatment of Cancer – Quality of Life Questionnaire Core 30
F	Fludarabine
FC	Fludarabine plus cyclophosphamide combination
HR	Hazard ratio
HRQoL	Health-related-quality-of life
ICER	Incremental cost effectiveness ratio
IgD	Immunoglobulin D
IgD	Immunoglobulin G
IgM	Immunoglobulin M
IRA	Initial response assessment
ITT	Intention to treat
i.v.	Intravenous
MRD	Minimal residual disease
NCI	National Cancer Institute
NCI-CTC	National Cancer Institute – common toxicity criteria
NHL	Non-Hodgkin's lymphoma
nPR	Nodular partial response
NE	Not evaluable
ORR	Overall response rate
OS	Overall survival

PD	Progressive disease
PFS	Progression-free survival
PR	Partial response
QALY	Quality adjusted life year
R-FC	Rituximab combined with fludaribine and cyclophosphamide
SAE	Serious adverse event
SD	Stable disease
SE	Standard error
TTNT	Time to new CLL treatment
TTP	Time to progression

1. Summary

1.1. Scope of the submission

The scope of the submission as determined by the National Institute for Health and Clinical Excellence (NICE) is to assess the use of rituximab in combination with fludarabine therapies versus fludarabine therapies alone or chlorambucil. This scope is somewhat narrower than that anticipated in the marketing authorisation and a type II variation to the European Medicines Evaluation Agency (EMEA) submitted by Roche (17th July 2008) for the licensing of rituximab in combination with any licensed chemotherapy regimens for the 1st line treatment of patients with CLL.

Within the submission the intervention assessed is:

- Rituximab in combination with fludarabine and cyclophosphamide (R-FC)

The comparators considered are:

- Fludarabine in combination with cyclophosphamide (FC)
- Chloramubucil monotherapy (C) (via a mixed treatment comparison model)

Other wider comparators, such as fludarabine alone, alemtuzum, and bendamustine (not outlined in the NICE scope for the appraisal, but included in the marketing authorisation submitted to the EMEA) are only considered further within the context of the mixed-treatment comparison model, linking evidence on the effects of rituximab in combination with fludarabine and cyclophosphamide to chlorambucil.

One subgroup of patients as outline in the scope issued by NICE is considered.

- Those with CLL who have p53 abnormalities

1.2. Summary of submitted clinical effectiveness evidence

The evidence for the submission is based on one phase III RCT comparing rituximab in combination with fludarabine and cyclophosphamide to fludarabine and cyclophosphamide alone for the 1st line treatment of people with CLL (the CLL-8 trial; n=810). Additional evidence is provided in the form of a mixed treatment comparison model linking results from the CLL-8 trial with other relevant comparators, notably chlorambucil; the comparison of which is included in the cost-utility model. Results showed statistically significant increased **progression-free survival (PFS)** with the rituximab combination therapy [median 39.8 months versus 32.2 months; HR 0.56 (95% CI: 0.43; 0.72)] compared with fludarabine and cyclophosphamide alone. An initial significant treatment benefit for the rituximab combination therapy compared to fludarabine and cyclophosphamide for **overall survival** was not maintained at a slightly longer follow-up time (median 25.4 months) [adjusted HR 0.72 (95% CI: 0.48; 1.09)]. **Response rates**, number of patients with **event-free survival**, and duration **of response** all favoured treatment with rituximab therapy.

The mixed treatment comparison (MTC) model indicated rituximab combination therapy to be significantly superior to chlorambucil alone for both disease progression and overall and complete response rates.

1.3. Summary of submitted cost-effectiveness evidence

Roche uses a Markov model, separating the disease process into the three states; Progression-Free Survival (PFS), Progressed, and Death, to analyse the cost-utility both of R-FC v. FC and R-FC v. chlorambucil. Effectiveness parameters for the model are derived from the CLL-8 trial data and a multiple treatment comparison is used to derive a hazard ratio value for R-FC v. chlorambucil. Costs are taken from established sources. Utilities values rely on estimates from a cited source which departs from the NICE reference case.

Estimated cost per QALY for the R-FC v. FC comparison in the base case is £13,189 and for R-FC v. chlorambucil comparison the base case is £6,422. A series of univariate (one-way), scenario, and probabilistic sensitivity analyses are provide which

show a strong probability these ICERs remain below normal accepted willingness-to-pay thresholds.

1.4. Commentary on the robustness of submitted evidence

1.4.1. Strengths

Clinical effectiveness

- The searches for clinical and cost-effectiveness data are appropriate and include all relevant studies.
- The identified RCT is well conducted and the findings likely to be reasonably robust.

Cost-effectiveness

- The approach taken to modelling is reasonable.
- The sources and justification of estimates were generally sound.

1.4.2. Weaknesses

- The evidence is based on only one completed and unpublished RCT.
- Sensitivity analysis is limited and does not fully investigate the uncertainty associated with differential values across arms, or with the structural assumptions of the model.
- Utility values are not drawn from an empirical study.

1.4.3. Areas of uncertainty

It is unclear whether the observed treatment benefit for use of rituximab combination therapy for PFS is associated with longer-term gains in overall survival and how plausible it is to extrapolate any PFS benefits in the longer term.

2. Background

2.1. Critique of manufacturer's description of underlying health problem

In section 2 of their submission, Roche report the incidence of chronic lymphocytic leukaemia in the United Kingdom based on a credible source. Section 4 then accurately outlines the common presentation, diagnosis and staging of the disease, with relevant guidelines as to whether patients should be considered for treatment presented. Relevant prognostic factors and the impact of genetic mutational status on prognosis are also discussed.

Overall, information provided about the underlying health problem is reasonably comprehensive and accurate.

2.2. Critique of manufacturer's overview of current service provision

The Roche submission indicates in section 4.1.5.1 that, historically, CLL has been managed with the aim of controlling the disease, minimising treatment-related toxicity and maximising quality-of-life. To this end, single agent chlorambucil, with acceptable efficacy together with only mild toxicity formed the main stay of 1st-line treatment therapies. However, since the publication of the UK LRF CCL-4 trial, which compared the use of combination FC therapy, fludarabine alone and chlorambucil,[1] there has been a move away from the use of chlorambucil, towards the use of the more aggressive combination FC treatment regimen with the aim of providing the best progression-free survival from 1st line treatment. This is reflected by figures on trends for the use of different therapies (chlorambucil, MabCampath, R-FC, and fludarabine alone) (Fig.1 p.25).[2] However, whilst it appears from these that the use of chlorambucil has declined steadily from Q4 2004 to Q2 2008, it appears that it is still used widely, in approximately 32% of patients, compared to FC which is used in 47%. Rituximab (in combination with other therapies) is currently used off-label in approximately 14% of patients. It would therefore appear to the ERG that although the

use of chlorambucil alone has declined in recent years, it is still widely used in current practice. This is a view supported by expert advice to the ERG.

Furthermore, a number of inconsistencies regarding the place for rituximab in 1st line CLL treatment advocated by Roche are noted. The submission highlights (p.25) that the initial management of patients with 17p deletions or other p52 abnormalities usually involves treatment with alemtuzumab, or enrolment into an ongoing 'high-risk' clinical trial but the submission subsequently states (p.28) "it is suggested that the initial treatment of any patient receiving treatment for CLL for the first time receives a combination therapy including rituximab, irrespective of age, performance status or genetic subgroup". The submission further states (p. 29) that "some clinicians will only reserve chlorambucil for the very, very frail and elderly and as a palliative measure, whereas some will consider it as their standard treatment, except for the very young and fit".

3. Critique of manufacturer's definition of decision problem

3.1. Population

The study population is defined in Roche's submission (Section 2) as people with previously untreated chronic lymphocytic leukaemia (CLL). This includes all patients who are considered eligible for 1st line treatment, as defined by The National Cancer Institute working group,[3] and is an appropriate description of the population under consideration.

One sub-group of patients is considered, those with p53 deletion at baseline. This sub-group is only considered further in relation to the outcome of PFS, and the sub-group is not specifically assessed in the cost-utility model. Given the scope of the appraisal to assess the effects of rituximab, fludarabine therapies and chlorambucil in this specific group, this limitation does not seem appropriate. Roche however state in their submission, that only a limited number of patients with p53 deletions (n=46) were included in CLL-8, and therefore the trial was not powered to detect differences in treatment effects in this patient sub-group.

3.2. Intervention

The intervention is rituximab (MabThera®). The evidence presented assesses the use of rituximab (R) in combination with fludarabine (F) and cyclophosphamide (C). Rituximab is presently used off license in combination with other therapies for the 1st-line treatment of CLL in the UK in approximately 14% of patients.[2] The scope of the appraisal as determined by the NICE[1] is narrower than that anticipated in the marketing authorisation and a type II variation to the European Medicines Evaluation Agency (EMEA) (submitted 17th July 2008) for the licensing of rituximab in combination with any licensed chemotherapy regimens for the 1st line treatment of patients with CLL. It was anticipated that opinion from the Committee on Medicinal

[1] Hereafter this will be referred to as 'the scope'

Products would follow on 18th December 2008, with full European Union marketing authorisation following 42 days after this. Therefore the estimated date for final authorisation is 30th January 2009.

The rituximab dose is 500 mg/m^2 IV body surface area for cycles 2-6, with a dose reduction to 375 mg/m^2 IV for cycle 1. Patients receiving rituximab in combination with fludarabine (25 mg/ m^2 i.v.) and cyclophosphamide (250 mg/m^2 i.v.) receive treatment on days one to three of each 28 day treatment cycle for a maximum of six cycles. Of note, the trial evidence presented used a fully iv combination regimen, but currently in the United Kingdom, 99% of F and C combination therapy is delivered orally.[2] Addition of rituximab to FC, therefore adds an i.v. component to what is usually a fully oral treatment regimen. The cost-utility model therefore assumes an oral FC regimen in both the intervention and comparator arms, with appropriate dose adjustments made to account for the lower bioavailability of the oral regimen.[4] This appears to be appropriate.

3.3. Comparators

The comparators are FC alone (with doses and delivery as stated above), and chlorambucil alone. The evidence for R-FC versus FC alone is based on a direct head-to-head comparison from the German CLL-8 trial.[5] Further supporting evidence on the effects of rituximab, in combination with either FC, pentostatin and C, or FC and mitoxantrone is provided from four phase II studies.[6-10]

A mixed treatment comparison (MTC) model is presented on the use of R-FC versus chlorambucil, alemtuxumab, fludarabine alone, and bendamustine. Chlorambucil had been included as a comparator in five identified trials,[1;11-14] and was administered at doses ranging from 20mg/m^2 to 70 mg/m^2 for between a range of six to 12 monthly cycles.

Only the comparators of FC and chlorambucil are considered further in the cost-utility model. This is in line with the appraisal scope as issued by NICE.

3.4. Outcomes

The clinical effectiveness outcomes considered are progression-free survival (PFS), overall survival (OS), event-free survival (EFS), disease-free survival (DFS), response rates, duration of response, time to new CLL treatment, health-related quality-of-life (HRQL) and adverse effects of treatment. These outcome measures are defined in Section 4.1.6.

The outcomes for the economic analysis are incremental cost per quality-adjusted life year (QALY), resource utilisation, and the cost of treating adverse events (blood transfusions and bone marrow transplant events). These outcomes are appropriate for assessing the impact of the different therapeutic regimens on both a range of clinical parameters, costs and QALYs.

3.5. Time Frame

The economic analysis took a lifetime time horizon, which was assumed to be 15 years. Advice to the ERG agrees that this is an appropriate time-frame, with the majority of CLL patients expected to be dead within this period.

3.6. Other relevant factors

Currently, the licensed, approved dose for rituximab in lymphoma (whether monotherapy or given in combination with chemotherapy) is 375mg/m^2. However, it has become clear from monotherapy dose finding studies in CLL that there is an increasing response in CLL patients as the dose is increased up to 2250mg/m^2.[15]

On the basis of this and further Phase II studies of R-FC in CLL,[6] a higher dose of 500mg/m^2 was decided upon as appropriate for CLL patients (with a dose reduction to 375mg/m^2 in cycle 1) in combination with FC, and used in the CLL-8 trial.[5]

The dosing of R-FC used in the CLL-8 trial is expected to be the approved dose in the Summary of Product Characteristics (SmPC).

4. Clinical effectiveness

4.1. Critique of manufacturer's approach

4.1.1. Description of manufacturers search strategy and comment on whether the search strategy was appropriate.

Searches were performed in the following databases from 1 January 2000 to date unless otherwise stated:

- EMBASE Dialog Datastar Search Date: 15 October 2008
- MEDLINE Dialog Datastar Search Date: 15 October 2008
- MEDLINE In-Process Dialog Datastar Search Date: 28 October 2008
- The Cochrane Library (all): Version: 2008 Issue 3. Search Date: 2 October 2008
- BIOSIS Dialog Datastar Search Date: 31 October 2008
- (BIOSIS for ASH annual meeting abstracts only)
- EMBASE ALERTS Dialog Datastar Search Date: 28 October 2008

Separate search strategies were provided for all databases by the manufacturer. The search strategies are inconsistent in the use of thesaurus and text word terms. In MEDLINE and EMBASE only thesaurus headings were used for population, intervention, and study design with the only text word being a qualifier of "chronic" in MEDLINE. The study design filter was exclusive in using only a thesaurus "publication type" filter. The ERG re-ran these searches with text words and a full clinical trials filter but no additional relevant trials were found. The Cochrane Library search used only text words terms in contrast to the MEDLINE and EMBASE

searches. The BIOSIS strategy used text words related to the population and intervention in the title only, while EMBASE ALERTS and MEDLINE-IN PROCESS (MEIP) used text words for population and intervention in all fields. The use of text words only in BIOSIS, EMBASE ALERTS, and MEIP is appropriate. No comparators or outcomes were specified to limit the searches in any of these databases. The only limits by study design were applied in MEDLINE and EMBASE.

The search strategy for the MTC results was provided later by the manufacturer. These searches were conducted in MEDLINE, EMBASE, and BIOSIS; however no information was provided on search dates or interface used. These searches include both thesaurus and text words and includes a far more inclusive study design type filter than the previous search strategy.

All the term combinations within the search strategies, as regards the CLL population and/or the intervention and resources used were appropriate, replicable, and the resulting hits appear correct related to the search date and database/interface used. However, there is a lack of text words in MEDLINE and EMBASE and a limited (exclusive) filter for study type. The MTC model search was complete, appropriate and more extensive although the dates of search and interface used are unclear.

4.1.2. Statement of the inclusion/exclusion criteria used in the study selection and comment on whether they were appropriate.

4.1.2.1. Direct head-to-head trials and relevant non-randomised controlled trials

In Roche's submission no explicit inclusion/exclusion criteria are stated for considering trials for inclusion in the systematic review of clinical effects. This is related to the immaturity of the available evidence base on the use of rituximab in combination with other therapeutic chemotherapy regimens for the 1st line treatment of patients with CLL; in which, to date, there are no published phase III comparative RCTs available. The inclusion/exclusion criteria are therefore implicitly driven by the inclusion criteria of the only relevant unpublished trial evidence to date; that from the German CLL-8 trial.[5] This trial:

- Included patients with previously untreated chronic lymphocytic leukaemia [Binet stage A (up to protocol amend number 1), Stage B and C disease] (for a description of the Binet stage classification criteria see Appendix 1).

- Compared the use of rituximab in combination with FC to FC alone.

- Assessed outcomes of: progression-free survival (PFS), overall survival (OS), event-free survival (EFS), disease-free survival (DFS), response rates, and duration of response, time to new CLL treatment, HRQL, and adverse effects of treatment.

In addition, evidence from four further phase II non-randomised controlled trials is included in the review of clinical effects to support the efficacy and tolerability of rituximab in combination with different chemotherapy regimens including FC, pentostatin and C, or FC and mitoxantrone.[6-10] the results of these studies are not pivotal to addressing the decision problem as specified in the scope, and therefore not considered further within the main report from the ERG. However, for completeness an outline of the key results from these studies is presented in Appendix 3.

These criteria seem appropriate to identify all relevant phase III RCTs and non-randomised phase II trials with a bearing on the clinical effects of rituximab. However, the submission does not explain the process used in study selection, for example how many people were involved in reviewing title/abstracts, how differences of opinion were resolved, and how the process of selection or rejection of retrieved papers was conducted.

4.1.2.2. Trials included in the mixed treatment comparison model (MTC)

The inclusion/exclusion criteria to identify studies for inclusion in the MTC were:

- Interventions: R-FC, chlorambucil, almetuzumab, fludarabine alone and bendamustine.

- Study design: randomised controlled trials (RCTs) that may be either be blinded or non-blinded, and published or unpublished.

- Study population: previously untreated patients requiring 1st-line treatment for CLL, and having a good ECOG performance status (0 to 2).

- Outcome measures: median HR for PFS, EFS, OS, disease-free survival, duration of response, time to new CLL treatment or death, and response rates.

Exclusion criteria were:

Trials that had assessed the use of cladribine, or high-dose chrlorambucil.

These inclusion criteria appear appropriate to identifying all relevant trials on the use of R-FC, chlorambucil, almetuzumab, fludarabine alone and bendamustine; thus allowing an indirect comparison on the relative treatment effects of R-FC and chlorambucil as appropriate to the initial scope specified by NICE. However, the submission again does not explain the process used in study selection or how many people were involved in this process.

4.1.3. Table of identified studies. What studies were included in the submission and what were excluded.

4.1.3.1. Direct head-to-head trials and relevant non-randomised controlled trials

The numbers in the search results presented by Roche tally and appear correct. Only one relevant unpublished study on the use of R-FC versus RF alone was identified; the German CLL-8 trial.[5] This is presented as the pivotal phase III randomised evidence on the use of R-FC versus FC alone. Data from four different sets of analyses of the trial are presented within the submission.

- The planned clinical interim analysis (cut-off July 3rd 2007; median follow-up 20.7 months) which became the main analysis as the study was halted at this time when it became clear that there was a treatment benefit for R-FC compared to FC for PFS.
- Snapshot analysis 1 (cut-off February 8th 2008; median follow-up 25.4 months).
- Snapshot analysis 2 (cut-off June 2008; median follow-up 25.5 months).
- Economic analyses snapshot (snapshot 4) (cut-off July 2008; median follow-up 26.4 months).

The interim (i.e. main) analysis forms the basis of the results presented,[5] with further results reported from both snapshot analyses 1 and 2. Data from a longer follow-up time (25.5 months) which was an analysis conducted independently from Roche by the lead trial investigators, Hallek and colleagues, has additionally been presented at the 2008 ASH conference, and for completeness is included in Table 1 of included relevant trials.[16]

Table 1 : Studies identified by Roche's search strategy

RCTs		Reference	Publication status
ML17102	Primary publication	Roche. CCL-8 Final Clinical Study Report[5]	unpublished trial report
	Up-dated data	Hallek and colleagues (2008)[16]	abstract
Phase II non-randomised controlled trials			
Regimen			
R-FC	Primary publication	Keating and colleagues (2005)[6]	full article
	Up-dated data	Tam and colleagues (2008)[7]	full article
R plus F (administered concurrently or sequentially)	Primary publication	Byrd and colleagues (2003)[8]	full article
Pentostatin in combination with R and C	Primary publication	Kay and colleagues (2007)[10]	full article
Mitoxantrone in combination with R-FC	Primary publication	Faderl and colleagues (2007)[9]	abstract

Data source: Roche submission section 6.2.1 to 6.2.5 (p. 32 – 37)

The details of one relevant on-going study in rituximab in combination with chlorambucil (Roche study MO20927) are also given. This was a single arm phase II study assessing the safety and efficacy of chlorambucil in combination with rituximab as initial treatment for CLL in patients ineligible for fludarabine-based treatments. It is anticipated that preliminary study results in the form of a conference abstract will be available in 2009.

4.1.3.2. Trials included in the mixed treatment comparison model (MTC)

The search results presented by Roche are inconsistent, and the numbers provided do not completely tally. According to the QUOROM flowchart (Roche submission section 6.6, Fig 9, p. 76)

- 683 hits were generated by initial searches (plus the unpublished CLL-8 trial was additionally identified; total of 684 hits)
- 668 were excluded on the basis of title and/or abstract
- 12 papers plus the CLL-8 trial (total n=13) were reviewed in full (implying that 684 –671 not 668 were excluded on the basis of title and/or abstract)
- 5 papers were excluded on the basis of perusal of the full text (exclusions sum correctly); leaving
- core included evidence base of 8 RCTs

Despite the fact that the numbers do not tally correctly, we do not believe that relevant studies have been missed. Details of the eight identified RCTs included in the MTC are presented below in Table 2.[1;5;11-14;17;18] Of these, five were included in the MTC for PFS giving comparisons between R-FC, FC, fludrabine alone, chlorambucil and alemtuxumab.[1;5;11;17;18]

Table 2 : Details of the trials included in the MTC

Ref	Interventions	Publication status
ML17102 (CLL-8)[5]	R-FC and FC	unpublished trial report
Catovsky and colleagues (2007)[1]	F, FC and chlorambucil	full article
Hillmen and colleagues (2007)[11]	chlorambucil and alemtuzumab	full article
Flinn and colleagues (2007)[17]	FC and F	full article
Rai and colleagues (2000)[12]	F and chlorambucil	full article
Eichhorst and colleagues (2006)[18]	FC and F	full article
Knauf and colleagues (2007)[13]	Bendamustine and chlorambucil	abstract
Eichhorst and colleagues (2007)[14]	F and chlorambucil	abstract

Data source: Roche submission section 6.6 (p. 73 – 76)

Five additional trials were excluded from the MTC.[19-23] These are listed below in Table 3 along with the reason stated by Roche for their exclusion. The ERG having assessed full paper copies of these trials agrees that their exclusion is appropriate.

Table 3 : Details of the trials excluded from the MTC

Ref	Reason stated by Roche for trial exclusion
Jaksic and colleagues (1997)[19]	The percentages within the Binet stages differed significantly across the arms. This implied that the estimated treatment effect would also be influenced by the difference in severity.
Jaksic and colleagues (2000)[20]	High-dose chlorambucil is not seen as relevant and does not indirectly link main comparators.
Karlsson and colleagues (2004)[21]	High-dose chlorambucil is not seen as relevant and does not indirectly link main comparators.
Leporrier and colleagues (2001)[22]	Definition of PFS was not defined as the time between randomisation and first time point of progression. The treatment scheme seemed to imply downwards bias: persons obtaining fludarabine at the start of the research period and not responding to the treatment at 3 months of treatment switch to cyclophosphamide, doxorubicin plus prednisone (CAP), where the CAP group was closed during the study because of toxicity. For the cyclophosphamide, vincristine, prednisone plus doxorubicin (CHOP) arm, a switch was only possible after 6 months of treatment. The fludarabine group therefore has a larger chance to obtain a less effective treatment (CAP) than the CHOP group, so that the results for the fludarabine group from time of randomisation seems to be biased downwards.
CLL Trialists' Collaboration Group (1999)[23]	Links chlorambucil with CHOP and CAP, which are both not relevant comparators, and no two-step path from chlorambucil to another relevant comparator goes via CHOP or CAP.

Data source: Roche submission section 6.6 (p. 74)

4.1.4. Details of any relevant studies that were not included in the submission?

As stated in section 4.1.1, the ERG re-ran the searches as they were specified in the submission, and no additional relevant trials were found. No further searches for studies on the clinical effects of rituximab in combination with any other chemotherapeutic regimen were undertaken by the ERG, but no further relevant studies were found from a review of the wider evidence, not included in the submission.

4.1.5. Description and critique of manufacturers approach to validity assessment

4.1.5.1. Direct head-to-head trials

The Roche submission provides a detailed narrative description of the methods used in the CLL-8 trial[5] (Section 6.3 p 37-59). The key points of their validity assessment are summarised along with the ERGs comments below in Table 4. The trial appears to be reasonably well designed, and orchestrated. Allocation concealment was adequate, however blinding (of both patients and outcome assessors) was not attained. This has the potential to introduce bias into the results, but it is not possible to know in which direction, if any, this bias would operate. The trial is randomised, and based upon adequate *a priori* sample size calculations in order to detect differences in treatment effect between the trial arms for the primary outcome measure of PFS. However, the trial was not powered to detect differences in treatment effect in patient sub-groups (most notably for the scope of this submission those with p53 deletion at baseline). Trial arms were well balanced in terms of known prognostic factors at baseline. However, the criteria for attaining a complete response (CR) was composed of both subjective and objective measurements, which could potentially bias outcome assessment for PFS. Again, however the direction in which this bias may operate is not known. Length of follow-up is adequate to detect differences in treatment effect between the R-FC and FC arms for PFS, and all effectiveness analyses were conducted on an intention to treat (ITT) basis. Adequate sensitivity analyses were conducted for all missing data. Safety analysis was conducted, as appropriate on a *per protocol* basis. Reliance on PFS as opposed to OS is the biggest limitation; but it has been noted that it may be hard to detect differences in OS. This is firstly due to the indolent nature of CLL, in which there is an expected survival time of five to 10 years. Therefore, within the time frame of a clinical trial, a reasonably small number of events (deaths) will occur. Data is therefore highly censored, and due to the small number of events there may not be sufficient power to detect differences in OS between study arms. Furthermore, when patients progress on one treatment, there is extensive cross-over to 2nd-line therapy, potentially leading to a confounding. This could potentially result in underestimation of any treatment benefit for R-FC compared to FC for OS.

Table 4 : Validity assessment of direct head-to-head trial evidence

Criterion	Roche response	ERG's comments
How was allocation concealed?	Patients were stratified according to country and disease stage before being randomised to trial treatment by a central procedure. Randomisation numbers were generated centrally by the German CLL Study Group and incorporated into a set of patient numbers and associated treatments that were sent to the investigators. The investigator entered the corresponding number for allocation to the treatment groups on each patients Clinical Report Form (CRF) (Section 2.4.3, p.40 full study trial report).[5]	This is appropriate and indicates that allocation concealment was attained.
What randomisation technique was used?	A block randomisation was used.	This is appropriate.
Was the sample size justified adequately?	The primary endpoint PFS was used to determine the sample size of the study based on data from the German CLL-4 trial.[14] The sample was based on 80% power to detect a HR of 0.74 for PFS between the two treatment arms at a two-sided α level of 5%.	This assumption is justified by the reported outcome PFS [HR: 0.56 (0.43; 0.72) indicating the trial was appropriately powered.
Has there been adequate follow-up?	Yes. Unblinding of the data and full analyses of all endpoints was mandated by the independent DSMB at the interim analysis given highly statistically significant differences between the two arms (R-FC and FC), making further follow-up very unlikely to change the magnitude of difference seen between the two arms. However the median PFS values in both arms may increase with time. It is appreciated that the median survival of CLL is between 5-10 years, and further results with a longer follow-up will become available.	Length of follow-up was appropriate. As noted in the report, analysis of all endpoints was mandated by the trial steering group at the time of the interim analysis due to significant differences between the R-FC and FC arms for PFS. As highlighted, the median survival for OS was not reached at the time of the interim analysis and further follow-up will continue.
Assessors aware of treatment allocation?	It is likely that assessors were aware of treatment allocation and outcomes in this study were assessed by the investigators. An independent assessment of the data was not performed. However the assessment of CLL post treatment is very objective and it is unlikely that this will have biased results.	Study investigators appear to be aware of treatment allocation, which could potentially bias outcome assessment. Complete response (CR) was assessed by a combination of objective haematological parameters and subjective assessments (physical examination results, absence of B-symptoms) and therefore could be subject to bias in its determination. This could potentially impact on the number of patients assessed as having PFS. OS assessment would not be affected by knowledge of allocation.

		There is potential for bias in the assessment of CR rates and therefore PFS. It is not known in which direction this bias may operate.
Was the design parallel group or cross-over ?	Parallel-group The primary end-point of PFS would not be influenced by post-study treatment, and start of a new (i.e. second line) CLL treatment post randomisation was not considered a reason for censoring.	Parallel-group design is appropriate for this type of trial. The primary endpoint of PFS was measured at appropriate time points, before cross-over to 2nd line treatment. Analysis of OS would be influenced by a change in 2nd-line treatment, and is likely to underestimate the potential OS benefit from PFS with R-FC.
Was the study carried out in UK and how does the population compare with patients who are likely to receive R-FC in The United Kingdom?	CLL-8 was an international study not including the UK. However, there are no obvious differences between the study population and non-trial patients requiring treatment for chronic lymphocytic leukaemia in the UK, except, perhaps that the study patients were slightly younger. Certainly the generally caucasian population in Germany (where over 500 of the 817 patients were recruited) would compare very favourably with a British population. Other countries involved in recruitment (Australia, Austria, Belgium, Czech Republic, Denmark, France, Italy, Spain, New Zealand and Israel), all provide a demographic of patients that would be very similar in general to the United Kingdom. Disproportionate recruitment of younger patients is a general problem in oncology clinical trials – the study had no upper age limit for participation, and the oldest patient recruited was 82.	The median age at onset in England Wales is between 65 and 70 . The trial contained patients with a median age at baseline of 61.0 (range: 30-81). However, 70% of patients were below the age of 65 years; 23% were ≤ 65 -≥ 70 years old, and 7% were older than 70 years. A disproportion number of trial participants were therefore younger than those commonly seen in practice. Patients with an ECOG performance status of 0-1 were included; (see Appendix 2 of this report for details of ECOG scale) implying that patients were reasonably fit and active. Trial participants may therefore have been healthier than those generally presenting in routine practice. At baseline, 5% of patients had Binet stage A disease; 64% had stage B disease and 31% had stage C disease (see Appendix 1 of the report for details of the Binet criteria) After protocol amendment 1; only patients with stage B or C disease were eligible for inclusion in the trial. However, the inclusion of 5% of patients with stage A disease, means this sub-group would potentially have a better prognosis, and do not represent patients who would be eligible for initial treatment in the UK. The sex ratio of male: female participants in the trial (74%; 26%) is representative of the sex ratio of patients presenting for 1st-line treatment in practice.
Was the dosage regimen acceptable and justifiable?	Currently, the licensed, approved dose for rituximab in lymphoma (whether monotherapy or given in combination with chemotherapy) is 375mg/m^2. It had become apparent	The dose regimen of rituximab was appropriate and based upon 2 dose finding studies.[6;7] The control arm in the RCT

	from monotherapy dose finding studies in CLL.[24] that there was an increasing response in CLL patients as the dose increased up to 2250mg/m^2. On the basis of this, groups starting Phase II studies of R-FC in CLL (specifically Keating and colleagues[6] at the MD Anderson Cancer Center in the USA felt that the lymphoma dose was not appropriate for CLL and a higher dose would be required. 500mg/m^2 was decided upon as an acceptable higher dose for CLL patients to use in combination with FC. The dosing chosen in CLL-8 was based on the MDAAC Phase II studies. A dose reduction of 375mg/m^2 in cycle 1 was chosen to minimise any potential cytokine release/ tumour lysis that may have been triggered by the known large circulating tumour burden in CLL. Thus the dosing of rituximab in this study was appropriate and consistent with Phase II R-FC studies in CLL. The dosing used in these Phase II studies was also rationalised and based upon a published dose-finding study as highlighted above. The dosing of fludarabine and cyclophosphamide was the same in both arms and based upon dosing that has been independently used in Phase II and III studies in CLL. The dosing of R-FC used in this study is expected to be the approved dose in the SmPC.	received the FC regimen in identical dosage to that provided as concomitant therapy in the F arm.
Were the inclusion and exclusion criteria appropriate?	The inclusion and exclusion criteria were appropriate and consistent with accepted and validated criteria for running CLL trials.	Trial inclusion criteria were appropriate; particularly after protocol amendment 1 which excluded patients with Binet stage A disease at baseline
Were both arms of the study group comparable?	Yes – as detailed in section 6.3.2, patient characteristics in both arms were well balanced at baseline.	The selection of baseline characteristics (Roche submission tables 7-10) seems reasonably comprehensive and the cohorts well matched.
Were appropriate statistical tests used?	Yes, fully detailed in section 6.3.5.	See section 4.2 below
Was an intention to treat analysis undertaken?	Yes, as fully detailed in section 6.3.5. Efficacy analyses and economic analysis are subsequently presented for the intention-to-treat population.	Intention to treat analysis were undertaken for all effectiveness outcomes. Safety data were based on per protocol analysis
Are there any confounding factors that may attenuate the	There are not thought to be any confounding factors that attenuate the interpretation of the primary	Agreed, and highlighted in the submission.

interpretation of the study ?	endpoint and most of the secondary endpoints. For the analysis of overall survival, it is likely that cross-over limits the ability to show an overall survival benefit in favour of R-FC, an issue that has been seen in a number of Phase III CLL studies.	

Data source: Roche submission section 6.3.6; table 13 (p. 54 – 56)

4.1.5.2. Mixed treatment comparison model

The validity of the eight trials included in the MTC were assessed by Roche using the 5-point Jadad scale.[25] The submission states (p. 77) that all the trials were randomised appropriately. It appears that only the trial by Hillmen and colleagues[11] was blinded, with blinding either being inadequate in the other seven trials or not reported.[1;5;12-14;17;18] No specific details are presented on whether patient withdrawals were reported in the trials, but it would appear from the overall total score attained in each of the trials that withdrawals were adequately reported in all trials, except that by Knauf and colleagues.[13] Overall, only one of the trials attained a maximum score on the scale of five,[11] six attained a score of three,[1;5;12;13;17;18] and one attained a score of 2.[13] A summary of the Jadad quality criteria indicators for each of the trials is presented in Table 5.

Table 5 : Validity assessment scores for trials included in the MTC

Reference	Jadad validity assessment criteria			Total score (out of 5)
	Randomisaton	Blinding	Patient attrition	
ML17102 (CLL-8)[5]	Y	N	Y	3
Catovsky and colleagues (2007)[1]	Y	N	Y	3
Hillmen and colleagues (2007)[26]	Y	Y	Y	5
Flinn and colleagues (2007)[17]	Y	N	Y	3
Rai and colleagues (2000)[12]	Y	N	Y	3
Eichhorst and colleagues (2006)[18]	Y	N	Y	3
Knauf and colleagues (2007)[13]	Y	N	N	2

Eichhorst and colleagues (2007)[14]	Y	N	Y	3

Data source: Roche submission section 6.3.6 (p. 77 - 79)

Use of the Jadad quality assessment criteria scale is open to criticism, for although it focuses on three important dimensions of internal validity, it potentially gives more weight to the quality of reporting than to actual methodological quality. For example, a statement on patient attrition will earn the point allocated to this domain, independently of how many patients were excluded or whether or not the data were analysed according to the ITT principle. Furthermore, the scale addresses the generation of allocation sequences, a domain not consistently related to bias,[27;28] but does not assess allocation concealment, which has been shown to be associated with exaggerated treatment effects.[27;28] The use of this scale to assess the validity of the trials included in the MTC is therefore a somewhat limited approach, which does not present a complete picture of the quality of the included trials.

4.1.6. Description and critique of manufacturers outcome selection

The outcomes on which Roche's submission focuses are a direct reflection of the outcomes assessed in the CLL-8 trial.[5]

- The primary effectiveness outcome was progression-free survival (PFS); defined as the time between randomisation and the date of the first documented disease progression (see below), relapse or death by any cause.

Secondary endpoints included:

- Overall survival (OS); defined as the time between randomisation and the date of death from any cause.

- Event-Free Survival (EFS); defined as the time between randomisation and the date of progressive disease, relapse, start of new CLL treatment or death by any cause.

- Disease-Free Survival (DFS); defined for all patients with a confirmed complete response (CR). This was calculated from the time of first documented CR to the documented relapsed or death from any cause.

- Duration of Response; defined for all patients who achieved any level of response [i.e. complete response (CR), nodular partial response (nPR), partial response (PR)]. This was calculated from the time of first documented response to treatment to the documented disease progression or death by any cause (see staging of response assessment below).

- Overall Response Rate; calculated as the sum of complete and partial responses.

- Time to New CLL Treatment (TTNT); calculated from the date of randomisation to the date of starting a new CLL treatment.

- Safety and tolerability were assessed by analysis of at least one grade 3 or 4 adverse event; serious adverse events, events leading to treatment discontinuation, dose modification or interruptions, or treatment related death.

- HRQL was measured using the Spitzer Quality of Life Index[29] every assessment visit and the European Organisation for Research and Treatment of Cancer-Quality of Life Questionnaire Core 30 (EORTC-QLQ-C30).[30] HRQL was assessed at baseline, after cycle 6 (initial staging), final staging and at months 6,12, 24, 36, 48, 60 and subsequent annual assessments in the follow-up period, regardless of progression or alternative therapy being initiated.

Quality of life data were not reported in the submission and are due to be presented in a future separate publication (Roche submission p.49)

Response Assessments

Evaluation of treatment outcome and disease progression was performed according to the standard criteria as defined by the National Cancer Institute (NCI) Sponsored Working Group on CLL as listed in Table 6 below.[3]

Table 6 : CLL response determination criteria

Outcome	Criteria for Classification
Complete response (CR)	Complete response (CR) required that a patient satisfied all of the following criteria for a period of at least 8 weeks: ▪ Absence of lymphadenopathy confirmed by physical examination and/or appropriate radiographic techniques (i.e. all lymph nodes ≤

	1cm in diameter) ▪ No hepatomegaly or splenomegaly by physical examination and/or appropriate radiographic techniques ▪ Absence of B-symptoms ▪ Normal blood count with: ➢ Lymphocytes <ULN ➢ Polymorphonuclear leukocytes ≥ 1.5 x 10^9/L ➢ Platelets > 100 x 10^9/L ➢ Hemoglobin > 11 g/dL ▪ Bone marrow aspirate and biopsy had to be performed 8 weeks after the clinical and laboratory results demonstrated that a CR was achieved. The marrow sample had to be normocellular for age with less than 30% lymphocytes. Lymphoid nodules had to be absent. If the bone marrow was hypocellular, a repeat biopsy was taken 4 weeks later and samples were re-reviewed in conjunction with the prior pathology.
Nodular partial response (nPR)	Some patients fulfilled all the criteria for a CR (listed above) but still had nodules of lymphocytes in the bone marrow despite a total lymphocyte proportion of < 30% . These patients therefore fulfilled the criteria of having attained a nodular partial response (nPR) and were listed as such in the trial report.
Partial Response (PR)	Partial response (PR) was present if patients demonstrated the following criteria for at least 8 weeks: ▪ Reduction in peripheral lymphocyte count by ≥50% from pre-treatment value ▪ Reduction in lymph node enlargement by ≥50% ▪ Reduction of hepato- and/or splenomegaly by ≥50% if enlarged at baseline Plus at least one of the following criteria: ▪ Polymorphonuclear leukocytes (granulocytes) ≥1.5x10^9/L or 50% improvement over baseline value ▪ Platelets > 100 X10^9/L or 50% improvement over baseline value ▪ Hemoglobin >11g/dL or 50% improvement over baseline value without blood transfusions
Progressive disease (PD)	Progressive disease (PD) was present if at least one of the following criteria was fulfilled: ≥ 50% increase in the sum of the products of the diameters of at least two lymph nodes (at least one node had to be ≥ 2 cm) or appearance of new

	lymph nodes or any new extra-nodal lesion (regardless of size) ≥ 50% increase in the size of the liver and/or spleen as determined by measurement below the relevant costal margin or by ultrasound/CT scan; appearance of palpable hepatomegaly or splenomegaly that was not previously present ≥ 50% increase in the absolute number of circulating lymphocytes to at least 5 x 10^9/L Transformation to a more aggressive histology (e.g. Richter's syndrome or prolymphocytic leukemia [PLL] with >55% of prolymphocytes) In cases of uncertain progression based on lymph node enlargement alone, measurements were repeated at least 2 weeks to exclude transient enlargement not indicative of progressive disease
Stable disease (DS)	Stable disease (SD) was considered to be present if the criteria for CR, nPR, PR and PD were not fulfilled

Data source: Roche submission section 6.3.4 (p. 49 - 51); based on the National Cancer Institute (NCI) Sponsored Working Group on CLL criteria.[3]

Interim staging

Interim staging of all patients was performed after 3 cycles of therapy, i.e. before starting cycle 4. All patients who showed at least a partial response (PR/CR) after the first 3 cycles continued treatment according to the protocol for a total of 6 cycles of therapy. Patients who showed no response (SD/PD) after the first 3 treatment cycles were withdrawn from study treatment and were eligible to receive an alternative regimen. An initial response assessment (IRA) was performed 4 weeks after the beginning of the last cycle of therapy, with final-staging performed at least 8 weeks after the IRA to confirm the response at IRA. Further follow-up examinations were performed every 3 months during years 1-3, every 6 months during years 4-5, and annually thereafter until year 8. It should be noted that response data differ between the times of the interim (i.e. main) analysis with a median follow-up time of 20.7 months[5] and that presented in the abstract by Hallek and colleagues with a median follow-up of 25.5 months.[16] This is due to the fact that the second analysis also included patients with a 'late CR'. This reflects the fact there were a number of patients who only fulfilled criteria of a PR at the end of treatment, and subsequently changed treatment status to a full CR or 'late CR' on further follow-up. This is a phenomenon commonly seen in CLL, and other indolent malignancies where it is well

accepted that the time of best response may not always be at the time of formal assessment, and an actual CR may only become apparent in the months following the last treatment.

Critique

This selection of outcomes is consistent with those specified in the NICE scope, and appears to provide a reasonable range of dimensions in which to assess the clinical effectiveness of R-FC, FC and chlorambucil. It is unfortunate that at the time of the submission, the HQOL data from the trial (Spitzer Quality of Life Index[29] and EORTC-QLQ-C30[30]) were not analysed and made available to the ERG. This would have provided reasonable randomised controlled trial evidence on the effects of R-FC versus FC on the impact of treatment from the patients' perspective.

The Roche submission (page 72) adequately highlights the limitations of OS as an outcome measure in CLL research; and the potential confounding effect of cross-over to second line treatment which commonly occurs in CLL trials, whereby showing an OS benefit in favour of any particular treatment is very difficult. It should be noted that, to date, no Phase III randomised controlled trials, including fludarabine in combination with cycholophospamide, fludarabine alone, and chlorambucil[1;14;17] for the 1st line treatment of patients with CLL have found an overall survival advantage in favour of any particular treatment arm. Some of these trials have shown trends towards OS benefits and it is well accepted that the phenomenon of cross-over is a major reason why showing an OS benefit is very difficult. Patients who are given the less efficacious treatment in one arm will tend to relapse earlier and often then be given the treatment that patients in the other arm received. In relation to the CLL-8 trial,[5] patients were not exclusively crossed-over between the treatment arms (R-FC to FC and vice versa) once they had progressive disease, but went on to receive a number of different 2nd or subsequent line treatment options. At the time of the interim (i.e. main) analysis with a median follow-up of 20.7 months, in the FC arm 35/59 (59%) patients who received subsequent therapy for CLL were known to have received rituximab, either in combination with a chemotherapy regimen or as a single agent, compared to 19/44 (43%) of patients in the R-FC arm.

4.2. Describe and critique the statistical approach used

4.2.1 Direct head-to-head trials

Intention to treat analyses was performed. Timing of assessments and the interim analysis were correctly determined *a priori*. All primary efficacy analyses on PFS, OS, EFS, DFS, and TTNT were properly analysed using Kaplan-Meier methods and non-stratified, two-sided log-rank tests. Additionally, stratified results (with stratification based on country and Binet stage at pre-therapeutic staging) were presented for PFS and OS. Response rates were compared using a two-sided Chi-square test. Sensitivity analyses were conducted imputing a 'worse case' scenario for all missing response data. For all efficacy outcome measures the significance level was set at 5%. No adjustment for the multiplicity of testing was performed for the secondary outcome measures.

All safety data calculations were performed for the per protocol population (i.e. all patients who received at least one course of study treatment). No formal significance testing of safety end-points was performed.

The statistical approach to the analyses of the trial data appears justified and as far as can be determined correct.

4.2.2. Mixed treatment comparison model

Trials in the MTC were analysed using a Bayesian MTC approach. For the outcome of PFS a Cox regression model was assumed, which implies the HR (relative efficacy with respect to instantaneous risk of an event) is constant over time. Log hazards for PFS were summarised across trials and interpreted in terms of medians, based on the PFS curve for R-FC presented in the CLL-8 trial report. For complete response and overall response, the odds ratio (OR) and relative risk (RR) were estimated.

For the CLL comparator treatments of interest, the results of individual trials were combined using both a Bayesian fixed and random effects model. Probabilities were calculated based on the posterior uncertainty distributions of the treatment effect relative to each of the treatments compared. Heterogeneity, and goodness of fit to the model data for the fixed effect model were compared with the credibility intervals of the goodness of fit test for the random effects model, based on the overlap of the

credibility intervals for the residual deviance. Results obtained using the MTC model method were validated by comparing them with results of using a conventional indirect comparison approach,[31] for the comparison between R-FC and chlomrabucil. Sensitivity analysis to examine differences in patient baseline Binet distribution and dose of chlorabucil,[1] were also undertaken.

The methods used to conduct the MTC appear appropriate and the results are therefore likely to be robust.

4.2.3. Summary Statement

The submission contains all the relevant studies and the relevant data within those studies. The submitted evidence adequately reflects the decision problem defined in the submission.

4.3. Summary of submitted evidence

4.3.1. Summary of results

4.3.1.1 Progression-free survival

Due to progression-free survival (PFS) being the primary outcome in the CLL-8 trial, this outcome is given prominence in the manufacturer's appraisal of clinical effectiveness (for details of how PFS was defined, see Section 4.1.6 above).

The primary results reproduced here are those reported at the time of the interim (i.e. main) analysis (clinical cut off July 4, 2007) with a median follow-up time of 20.7 months. For completeness the results for PFS at the times of snapshot analysis 1 (median follow-up time of 25.4 months) and snapshot analysis 2 (median follow-up time 25.5 months) are also reported in Table 7.

Table 7 : Summary of Progression-Free Survival

Trial: CLL-8[5]		
Interim analysis (median follow-up: 20.7 months)		
	FC	**R-FC**
N	407	403
Median PFS – days (95% CI)	981.0 (935; 1069)	1212.0 (1098; 1400)
p-value (Log-rank Test)	p<0.0001	
Unstratified HR (adjusted) (95% CI)	0.56 (0.43; 0.72)	
p-value (Wald Test)	p<0.0001	
Stratified HR (unadjusted) (95% CI)	0.53 (0.41; 0.68)	
p-value (Wald Test)	p<0.0001	
Snapshot analysis 1 (median follow-up: 25.4 months)		
	FC	R-FC
N	407	403
Median PFS – days (95% CI)	988.0 (846; 1086)	1303.0 (1156; 1400)
p-value (Log-rank Test)	p<0.0001	
HR (adjusted) (95% CI)	0.60 (0.48; 0.76)	
p-value (Wald Test)	p<0.0001	
Snapshot analysis 2 (median follow-up: 25.5 months)		
	FC	R-FC
N	N evaluable = 787 (not reported separately by trial arm)	
% PFS	62.3%	76.6%
HR	0.59	
p-value	p<0.0001	

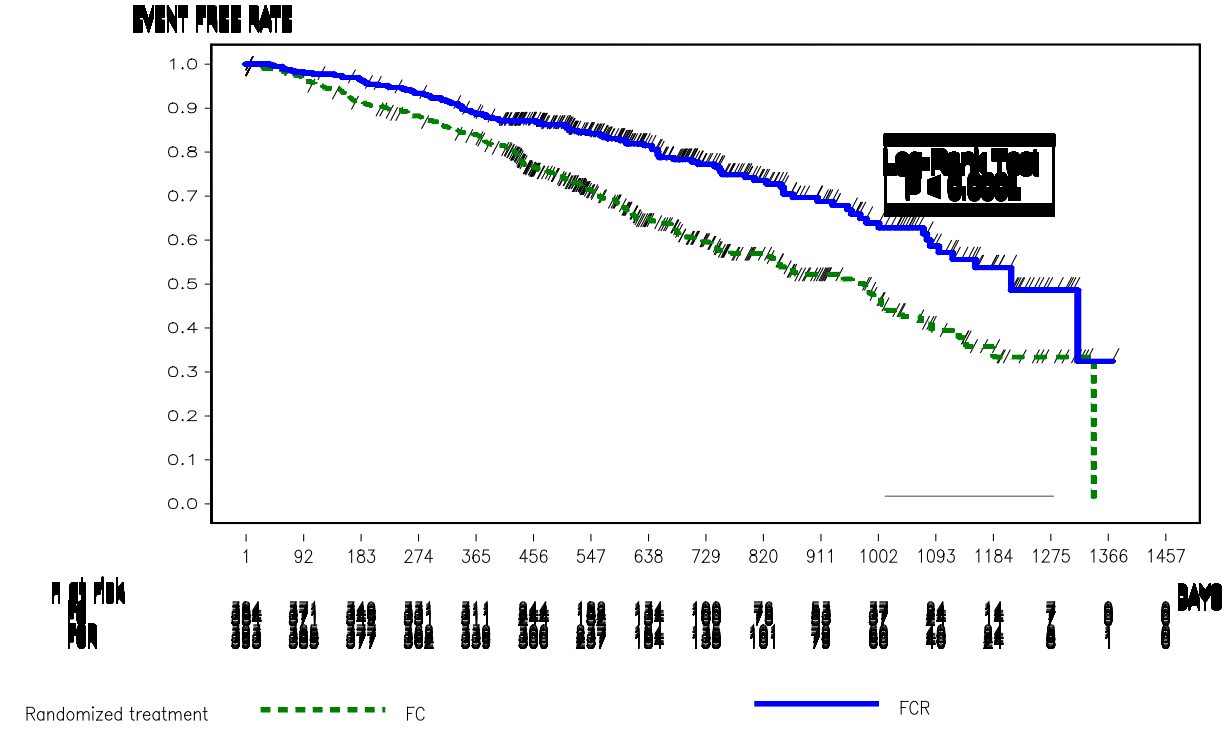
At the time of the interim (i.e. main) analysis a total of 254 patients (31%; 152 patients in FC, 102 patients in R-FC arm) had progressed (127 patients in FC, 85 patients in R-FC) or died (25 patients in FC, 17 patients in R-FC). In the FC arm, 37% (152/407) of patients had experienced an event compared to 25% (102/403) in the R-FC arm. R-FC significantly prolonged the median PFS when compared to the FC regimen alone ($p<0.0001$; log-rank test). The Kaplan-Meier estimated median PFS was 39.8 months with R-FC compared with 32.2 months with FC alone. The risk of having a PFS event (progression or death, whichever occurred first) was statistically significantly reduced by 44% (adjusted Hazard Ratio [HR] 0.56; 95% CI [0.43; 0.72]; $p < 0.0001$, Wald test) for patients in the R-FC compared to the FC arm. Seventy-seven percent of the patients in the R-FC arm, and 60% of those in the FC arm, were progression-free at two years. (Submission p. 60)

Of note, the results of the stratified analysis (with stratification based on country and Binet stage at pre-therapeutic staging) of PFS were similar to the non-stratified analysis, as depicted in table 7. Additionally, the results of the Kaplan-Meier estimates of PFS at snapshot analysis 1 (median follow-up: 25.4 months) were consistent with those of the interim (i.e. main) analysis. At this time point a total of 296 patients (171 patients on FC, 124 patients on R-FC) had died or progressed; an approximate 5% increase in PFS or deaths compared to the original analysis. Median PFS was significantly longer for patients in the R-FC arm (42.8 months [1302 days]) than for those in the FC arm (32.5 months [988 days]); $p<0.001$ log-rank test. This consistency of results for PFS between the interim and longer-term follow-up indicates a more robust PFS advantage for patients treated with R-FC compared to FC, as more patients had experienced an event at longer follow-up and therefore the results are less dependent on censoring (Submission p. 70).

Kaplan-Meier curves for PFS at the time of the interim analysis (20.7 months follow-up) are presented in Figure 1.

Figure 1 : Kaplan-Meier Curve of Progression-Free Survival at the time of the interim analysis

eg_pfskm_I Kaplan-Meier Plot Of Progression-Free Survival (Censored Observations Shown)
Protocol(s): ML17102 (I17102G)
Analysis Population: Intent-To-Treat Population (N=810)
Snapshot Date: 08FEB2008 Cutoff Date: 04JUL2007



Program : $PROD/cd11899a/i17102a/eg_pfskm.sas / Output : $PROD/cd11899a/i17102g/reports/eg_pfskm_I.cgm
18MAR2008 20:41

Data source: Roche submission section 6.4.2; figure 6 (p. 61)

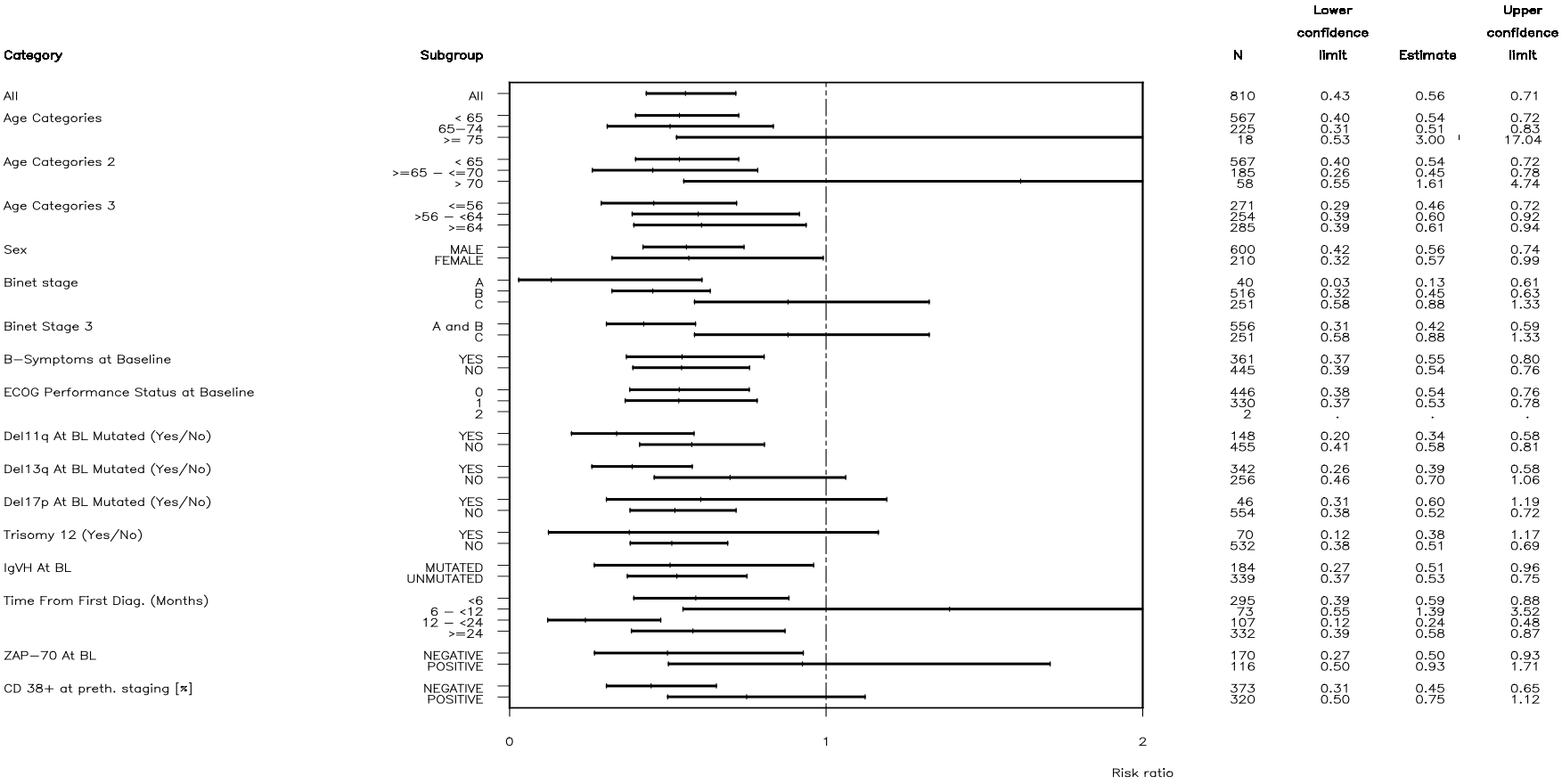
PFS in sub-groups

Roche's submission (pages 68 and 69) provides details of hazard ratios for PFS by sub-groups according to a variety of different criteria pertaining to the participants' characteristics (age, sex), clinical history (Binet stage at baseline, B-symptoms at baseline, ECOG performance status at baseline, time from first diagnosis) and prognostic biomarkers (del 11q mutated, del 13q mutated, del 17p mutated, trisomy 12, IgVH, ZAP-70+ at baseline). The submission adequately highlights that the trial was not powered to detect differences in treatment effect between sub-groups, and therefore any results should be interpreted in light of this. The forest plot of the Hazard Ratios for PFS by sub-groups at the time of the interim analysis is shown in Figure 2.

Figure 2 : Forest plot of Hazard Ratios for Progression-free survival by sub-groups

eg_pfscox_hr1_all_I Hazard Ratios And 95%-Confidence Intervals For PFS

Protocol(s): ML17102 (I17102G)
Analysis Population: Intent-To-Treat Population (N=810)
Snapshot Date: 08FEB2008 Cutoff Date: 04JUL2007



Data source: Roche submission section 6.4.2; figure 8 (p. 68)

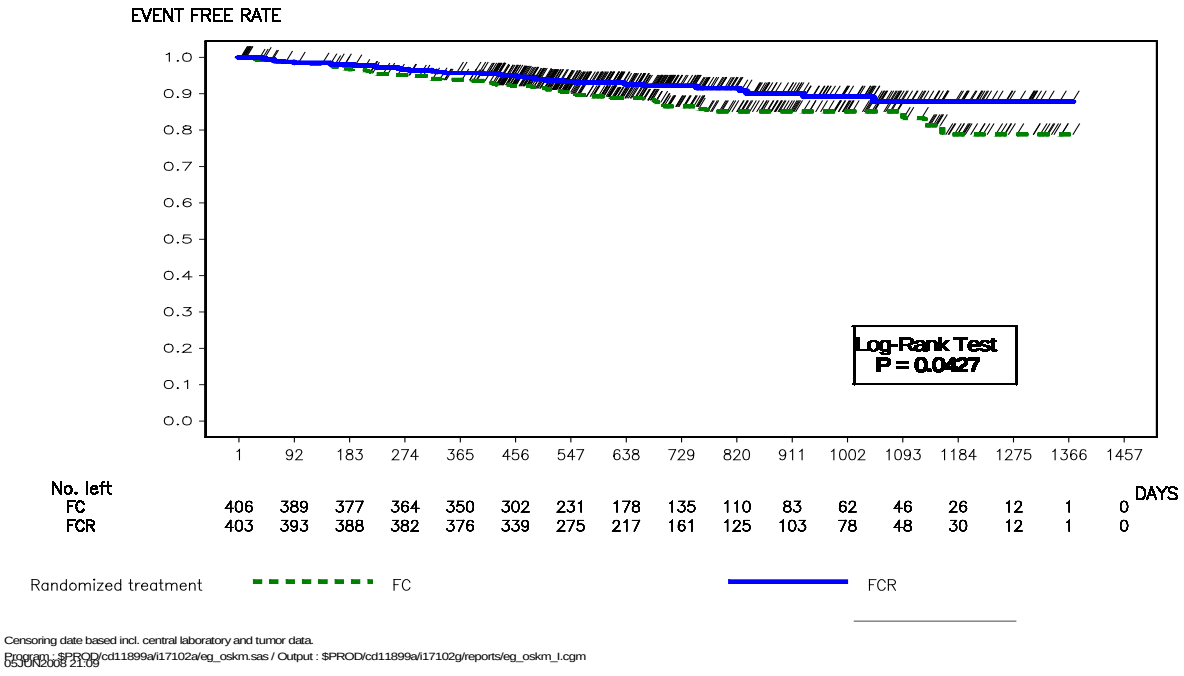
A significant PFS benefit for those taking R-FC compared to FC alone was observed in the majority of sub-groups for which data were presented. It is noticeable that in all sub-groups analysed according to the stratification factor of Binet stage at trial entry, the risk of PFS or death was decreased in the R-FC arm compared to FC alone. However, this risk reduction was more pronounced in the group of patients with stage A disease (unadjusted HR 0.13, 95% CI [0.03; 0.61]; p = 0.0093) and stage B disease (unadjusted HR 0.46, 95% CI [0.32; 0.63]; p <0.0001). In the patient subgroup with stage C disease, whilst the point estimates of the HR favours treatment with R-FC compared to FC alone (unadjusted HR 0.88, 95% CI [0.58; 1.33]; p = 0.5406); the bounds of the upper 95% CI crosses 1 denoting some uncertainty on the reality of any treatment benefit for R-FC compared to FC alone in this patient sub-group.

Only 46 patients p53 abnormalities (del 17p mutated) at baseline were included in the trial. The point estimates for the HR (unadjusted HR 0.6, 95% CI [0.31; 1.19] favoured a treatment benefit with R-FC compared with FC alone, but bounds of the upper CI again cross 1. Therefore there is some uncertainty regarding any potential treatment benefit in this traditionally more challenging to manage sub-group of patients for R-FC as compared with FC alone.

4.3.1.2 Overall survival

At the time of the interim analysis a total of 81 patients had died; 33 patients (8.2%) in the R-FC and 48 patients (11.8%) in the FC arm. Therefore data at this time were highly censored and follow-up was far from complete. Overall survival was significantly improved in the R-FC arm compared to the FC arm (p=0.0427, log-rank test), and reduced the risk of death by 36% (adjusted HR 0.64, 95% CI [0.41; 1.00]; p=0.0487, Wald-test). However, due to the immaturity of follow-up and the number of events that had occurred, the median survival time could not be estimated in both treatment arms. The Kaplan-Meier estimated 24-month survival rates as shown in Figure 3 were 92% in the R-FC arm and 87% in the FC arm. The results of the stratified analysis (factoring in country and Binet stage at pre-therapeutic staging) for OS were consistent with those of the non-stratified analysis (unadjusted HR 0.60, 95% CI [0.38; 0.94]; p=0.0250, Wald-test).

Figure 3 : Kaplan Meier Curve of Overall Survival



Data source: Roche submission section 6.4.2; figure 7 (p. 63)

With an additional 4.8 months of observation [(snapshot analysis 1 (median follow-up: 25.4 months)] there was no longer any OS advantage for treatment with R-FC compared with FC (adjusted HR 0.72: 95% CI [0.48; 1.09], p = 0.1252, Wald test). At this time there was a non-statistically significant trend observed in favour of treatment with R-FC with a 28% reduction in the risk of death. However, the data at this time point remained highly censored with 88% of patients in the FC arm (95% CI [85%, 92%]) and 91% of patients in the R-FC arm (95% CI [98%, 95%] still alive, and it is not possible to tell whether any small differences in OS between the two treatment arms at this relatively short follow-up period could reliably be extrapolated across the generally anticipated five to 10 year time horizon of the disease course.

4.3.1.3 Event-free survival

In the R-FC arm, 26.3% (106/403) of patients experienced an EFS event (disease progression, relapse, death or start of a new CLL treatment) compared to 39.3% (160/407) patients in the FC arm. Most of the events reported in both trial arms were disease progressions (77 events in R-FC; 116 events in FC); therefore any differences in EFS between the trial arms is driven by differences in the rates of disease

progression. Similarly, differences in the number of patients who received a new treatment (3.5% in R-FC; 5.9% FC) before reporting progression or death) also reflects differences in the number of patients with an insufficient response (stable disease) between the treatment groups. (Submission p. 63 and 64)

The median EFS was significantly longer in the R-FC arm, (39.8 months) compared to the FC arm (31.1 months); ($p<0.0001$, log-rank test). There was a statistically significant decrease of the risk of experiencing an EFS event of 45% (adjusted HR 0.55; 95% CI [0.43; 0.70]; $p<0.0001$, Wald test) in this trial arm compared to the FC arm. At 2 years, 75% of the patients in the R-FC arm, and 57% of those in the FC arm, were event-free.

4.3.1.4. Response rates

Response to therapy was categorised according to the standard criteria defined by the National Cancer Institute (NCI) Sponsored Working Group on CLL[5] (see Section 4.1.6).

Table 8 below shows the response rates at both (a) the time of the interim analysis (median 20.3 months) and (b) those for further follow-up at a median of 25.5 months[16]

Table 8: Summary of End-of-Treatment Response Rate

Trial: CLL-8[5]	**FC**	**R-FC**	**Difference in response rates (95% CI)**	**p-value**
Interim analysis (median follow-up: 20.7 months)				
N	407	403	-	-
Overall Response Rate (95% CI)	296 (72.7%)	347 (86.1%)	13.38 [7.8;19.0]	<0.0001
Complete Response Rate (CR) (95% CI)	170 (17.2%) [13.7; 21.2]	45 (36.0%) [31.3;40.9]	18.78 [12.7;24.9]	<0.0001
Partial Response Rage /nPR (95% CI)	226 (55.5%) [50.6; 60.4]	202 (50.1%) [45.1;55.1]	-5.40 [-12.4;1.6]	0.1234

Stable Disease (95% CI)	31 (7.6%) [5.2; 10.6]	19 (4.7%) [2.9;7.3]	NR	NR
Progressive Disease (95% CI)	31 (7.6%) [5.2; 10.6]	14 (3.5%) [1.9;5.8]	NR	NR
Missing (No response assessment) (95% CI)	49 (12.0%)	23 (5.7%)	-	-
Snapshot 2 (median follow-up: 25.5 months)[16]				
	FC	**R-FC**		
N	N evaluable = 761 (not reported separately by trial arm)		-	-
Overall Response Rate	328 (88%)	370 (95%)	-	p=0.001
Complete Response Rate	27%	52%	-	p<0.0001

Data source: Roche submission section 6.4.2; table 17 (p. 64) and table 20 (p. 71)

At the time of the interim (i.e. main) analysis, the proportion of patients with an objective response (CR/PR) was significantly higher in the R-FC arm (86.1%; 347/403) compared to the FC arm (72.7%; 296/407) (p<0.0001, Chi-square test). The complete response rate was doubled in the R-FC arm (145/403; 36.0%) compared to the FC arm (70/407; 17.2%) (p<0.0001, Chi-square test). There were more patients with stable disease (31/407; 7.6% in FC versus 19/403; 4.7% in R-FC) or progressive disease (31/407; 7.6% in FC versus 14/403; 3.5% in R-FC) in the FC arm compared to the R-FC arm.

Of note, there were more patients in both the R-FC arm and FC arm with either an overall response or complete response at the time snapshot analysis 2 compared to the time of the interim analysis.[16] This is due, as previously mentioned in section 4.1.6, to patients with a 'late CR' being included in this further analysis, whereas they would have been classified as having a PR at a shorter post-treatment follow-up time.

4.3.1.5. Duration of response rates

Duration of response was assessed in patients who had a confirmed response (CR, nPR or PR). Eighty percent (80%) of the patients in the R-FC arm, and 69% of those in the

FC arm, were event-free at two years. The median duration of response was significantly longer in the R-FC arm (40.2 months) than in the FC arm (34.7 months) (p= 0.004, log-rank test) and the adjusted HR was 0.61 (95% CI: 0.43; 0.85; p= 0.0036 Wald test) in the R-FC arm (Submission p. 66)

4.3.1.6. Disease-free survival

Disease-free survival (DFS) was defined for patients with a confirmed CR only. Ninety-one patients in the FC arm (91/407 patients, 22%) and 186 patients (186/403 patients, 46%) in the R-FC arm were included in the interim analysis. This analysis also included patients with a 'late response' of CR. At the time of analysis (median follow-up: 20.7 months), 12.1% (11/91 patients) of the patients in the FC arm experienced an event compared to 12.4% (23/186 patients) in the R-FC arm. The median was not reached in either arm and the number of events in both arms was very small (reflecting the generally favourable outlook for patients who achieved a CR) (Submission p. 66)

4.3.1.7. Time to new CLL treatment

At the time of the interim analysis, a total of 157 patients (22.1% [90/407patients] in the FC arm and 16.6% [67/403 patients]) in R-FC arm had started a new treatment for CLL (59 patients in the FC arm, and 44 patients in the R-FC) or had died (31 patients in the FC arm and 23 patients in the R-FC arm). The median time to new CLL treatment or death could not be estimated for both arms, however, the difference between arms was statistically significant (p=0.0052). The risk of receiving a new CLL therapy or of death was reduced by 35% in the R-FC arm compared to the FC alone regimen [(adjusted HR =0.65 (95% CI: 0.47; 0.90)]; p = 0.0082 (Wald-Test). At 2 years, 74% of patients in the FC arm had not received a new anti-CLL therapy or died, compared to 85% of patients in the R-FC arm (Submission p. 66)

Summary of clinical effectiveness outcomes from CLL-8

A summary of the key clinical effectiveness results from CLL-8 is shown in Table 8. This shows the significant treatment benefit for R-FC compared with FC for PFS, complete response rates, duration of response, event-free survival, and time to new CLL treatment. There were no significant differences between the comparators in terms of the number of patients achieving a partial response. It also highlights that the initial treatment benefit for the R-FC regimen in terms of OS observed at the time of

the interim analysis (median follow-up time 20.7 months) was no longer significantly different from that in the FC arm at the slightly longer follow-up.

Table 8 : Summary results for CLL-8

Outcomes	FC (n=407)					R-FC (n=403)					
	Analysis	**N with events**	**%**	**Median time in months**	**95% CI**	**N with events**	**%**	**Median time in months**	**95% CI**	**Hazard ratio 95% CI**	p-value
Progression-free survival											
Number with progression	Interim	152	37.3%	32.2	-	102	25.3%	39.8	-	0.56 (0.43; 0.72)	p<0.0001
	Snapshot 1	171	42.0%	32.5	-	125	31.0%	42.8	-	0.60 (0.48; 0.76)	p<0.0001
Overall survival											
	Interim	48	11.8%	NR	-	33	8.2%	NR	-	0.64 (0.41; 1.00)	p=0.048
	Snapshot 1	52	12.8%	NR	-	42	10.4%	NR	-	0.72 (0.42; 1.09)	p=0.1252
Event-free survival											
	Interim	160	39.3%	31.1	-	106	16.3%	39.8	-	0.55 (0.43; 0.70)	p<0.0001
Response rate											
CR	Interim	70	17.2%	-	13.7; 21.1	145	36.0%	-	31.3; 40.9	-	p<0.0001

PR		226	55.5%	-	50.6; 60.4	202	50.1%	-	45.1; 55.1	-	p=0.1234
SD		31	7.6%	-	5.2; 10.6	19	4.7%	-	2.9; 7.3	-	-
PD		31	7.6%	-	5.2; 10.6	14	3.5%	-	1.9; 5.8	-	-
CR/PR		296	72.7%	-	68.1; 77.0	347	86.1%	-	82.3; 89.3	-	p<0.0001
Disease free survival [a]											
	FC (n=91)					**R-FC (n=186)**					
	Interim	11	12.1%	NR	-	23	12.4%	NR	-	-	-
Duration of response											
	Interim	-	-	34.7	-	-	-	40.2	-	0.61 (0.43; 0.85)	p=0.0036
Time to new CLL treatment											
	Interim	90	22.1%	NR	-	67	16.6%	NR	-	0.65 (0.47; 0.90)	p=0.0052

Key: [a] Disease-free survival was only assessed in patients with CR (FC, n=91 and R-FC, n=186). This is used as the denominator for % reported; - data not reported; NR median time not yet reached; Data source: Roche submission section 6.4.2; tables 14 and 17 (p. 58 – 66)

4.3.1.8 Adverse events

In CLL-8 safety was analysed in the *per protocol* (PP) population, which consisted of all patients who received at least one dose of study medication. In the R-FC arm the PP population included 397/404 randomised patients and in the FC arm 396/407 randomised patients. The extent of study drug exposure in both arms is displayed in Table 9.

Table 9 : Number of treatment cycles received in the R-FC and FC treatment arms

Patients receiving at least x cycles	FC (n=396); No (%)	R-FC (n=397)	ALL (n=793)
1	396 (100.0%	397 (100.0%)	793 (100.0%)
2	366 (92.4%)	384 (96.7%)	750 (94.6%)
3	342 (86.4%)	364 (91.7%)	706 (89.0%)
4	313 (79.0%)	342 (86.1%)	655 (82.6%)
5	289 (73.0%)	317 (79.8%)	606 (76.4%)
6	273 (68.9%)	299 (75.3%)	572 (72.1%)

Key: x corresponds to the number of cycles received; Data source: Roche submission section 6.7.2; table 30 (p. 84)

Of those patients treated with R-FC, 75% received all scheduled cycles of therapy compared to 69% of those treated with FC. During the follow-up stage patients were prematurely withdrawn in both arms. A total of 47 patients (18 R-FC; 29 FC) were withdrawn due to safety concerns, including death, and 213 (114 R-FC; 89 FC) were withdrawn due to non-safety reasons, primarily insufficient response/progressive disease (70 R-FC; 98 FC). As the Roche submission highlights (p. 47), the difference between the arms in the number of patients receiving all scheduled treatment cycles was primarily due to the higher number of patients in the FC arm with insufficient responses at interim staging or withdrawals from the trial for administrative reasons.

In CLL-8 only grade 3 or 4 adverse events (AEs) or serious adverse events (SAEs) were collected. Serious adverse events were those defined as meeting the criteria of being fatal; life-threatening; disabling; requiring hospitalisation or prolongation of existing hospitalisation; being a congenital anomaly; or being a medically significant event(s) requiring intervention to prevent the previous outcomes. Lab abnormalities were only considered AEs if they were accompanied by clinical symptoms, led to a

change in study medication or required a change in concomitant therapy. All AEs were graded according to the NCI-CTC Version 2.0. Anemia and thrombocytopenia were graded according to the criteria for patients in leukemia studies. These definitions are widely used in oncology trials and are appropriate.

At the time of the interim analysis with a follow-up of 20.7 months the incidence of grade 3 or 4 AEs and SAEs was higher in the R-FC arm, while the number of all deaths was higher in the FC arm. AEs leading to dose modifications were more frequent in the R-FC arm than the FC arm. However, AEs leading to treatment discontinuation occurred with the same frequency in both arms (18%). There were no differences in the rate of deaths considered related to therapy (2%). Table 10 gives an overview of the adverse events in the trial.

Table 10 : Overview of adverse events in the CLL-8 trial

	Number of Patients (%)	
	FC (n=396)	**R-FC (n=397)**
Grade 3 or 4 AE	246 (62%)	304 (77%)
Serious AE	162 (41%)	182 (46%)
AE leading to treatment discontinuation	70 (18%)	71 (18%)
AE leading to dose modification/interruption	80 (20%)	133 (34%)
Treatment-related death	8 (2%)	6 (2%)

Data source: Roche submission section 6.7.2.2; (p. 85)

As can be seen from the table, the number of patients reporting at least one grade 3 or 4 AE was higher in the R-FC arm (77%) compared to the FC arm (62%). This was mostly due to a higher incidence of blood and lymphatic system disorders (57% R-FC versus 41% FC), which were mostly neutropenia and leucopenia. This is highlighted in the differences in the grade 3 or 4 AEs that occurred with at least 2% higher incidence between the R-FC and FC treatment arms (see Table 11).

Table 11 : Grade 3 or 4 AEs that occurred with an at least 2% higher incidence in one of the treatment arms

	FC (%)	R-FC (%)
Higher incidence in the R-FC arm compared to the FC arm		
Neutropenia	30%	19%
Leucopenia	23%	12%
Febrile neutropenia	9%	6%
Pancytopenia	3%	1%
Higher incidence in the R-FC arm compared to the FC arm		
Thrombocytopenia	10%	7%
Anaemia	7%	4%
Pyrexia	5%	3%

A summary of all grade 3 or 4 infections and infestations, and blood and lymphatic system disorders that occurred with an incidence of at least 1% in either of the treatment arms is given in Table 12.

Table 12 : summary of the grade 3 or 4 infections and infestations that occurred with an incidence of at least 1%[a]

Infections and infestations	FC (n=396) (%)	R-FC (n=397) (%)
Pneumonia	19 (5%)	14 (4%)
Herpes Zoster	12 (3%)	9 (2%)
Sepsis	8 (2%)	5 (1%)
Bronchitis	6 (2%)	5 (1%)
Infection	3 (1%)	6 (2%)
Sinusitis	3 (1%)	4 (1%)

Neutropenic infection	-	4 (1%)
Blood and lymphatic system disorders		
Neutropenia	75 (19%)	119 (30%)
Leukopenia	46 (12%)	93 (23%)
Thrombocytopenia	39 (10%)	26 (7%)
Febrile neutropania	22 (6%)	37 (9%)
Anaemia	26 (7%)	16 (4%)
Pancytopenia	5 (1%)	13 (3%)
Lymphopenia	6 (2%)	7 (2%)
Granulocytopenia	5 (1%)	7 (2%)
[a] most of the grade 3 or 4 infections in both arms were considered a SAE		

Deaths

A total of 80/793 patients (10%) in the safety population had died at the time of the main analysis. There were more deaths in the FC arm than in the R-FC arm (12% versus 8%).

The most common causes of death were infections (5% in FC versus 3% in R-FC) including sepsis, and neoplasms (including death due to PD; 4% in FC versus 3% in R-FC). Approximately 1% of patients in each treatment arm died due to a cardiac disorder.

In 8 FC patients (2%) and 6 R-FC patients (2%), the investigator judged the death to be related to study treatment.

Serious Adverse Events

The incidence of SAEs was 46% (182) patients in the R-FC compared to 41% (162 patients) in the FC arm. Serious infections and serious blood and lymphatic system disorders were most commonly reported in both treatment arms. The incidence for serious infections was 18% in R-FC versus 15% in FC; the incidence of serious blood

and lymphatic system disorders was 6% higher in the rituximab arm (17% in R-FC versus 11% in FC). All other SAEs were balanced between the treatment arms. Approximately half of the SAEs were considered by the investigators as related to treatment.

Adverse Events Leading to Treatment Discontinuation

The proportions of patients who discontinued study treatment due to AEs were equal in the two treatment arms; 18% in R-FC versus 18% in FC. Consistent with the overall pattern of AEs, the most common AEs that led to withdrawal were blood and lymphatic system disorders (12% R-FC versus 10% FC) and infections/infestations (2% in both treatment arms respectively).

Adverse events occurring on the first- or second-day of cycle treatment

AEs with a date of onset on the day of, or the day after, the start of any cycle treatment were analysed separately in order to describe the profile of potential infusion-related events. In the R-FC arm 16% of patients compared to 9% of patients in the FC arm had a grade 3 or 4 AE on the day of or the next day after start of cycle treatment. The most common events reported were blood and lymphatic system disorders (5% and 4% for R-FC and FC respectively), general disorders and administration site conditions (2% R-FC and <1% FC).

However it should however be highlighted that the trial used a completely iv dosing regimen, unlike that used in the UK in which FC is generally delivered orally, whilst R is delivered iv. To what extent this will potentially limit infusion related adverse events however is unknown.

Tumour Lysis Syndrome

A higher incidence of tumour lysis syndrome (TLS) was observed in the FC arm compared to the R-FC arm (9 patients FC versus 3 patients R-FC). Almost all of these events were of grade 3 and 4 intensity. Five of these in the FC arm and 2 events in the R-FC arm were classified as SAE, none of them were fatal.

Progression-Free Survival

Of the eight identified studies, five reported hazard ratios with respect to the primary outcome (PFS) and were therefore included in the MTC model.[1;5;11;17;18] allowing for a comparison of results from CLL-8. A list of these trials, the comparators considered

and the corresponding HR between treatment arms in each of the trials is shown in Table 13.

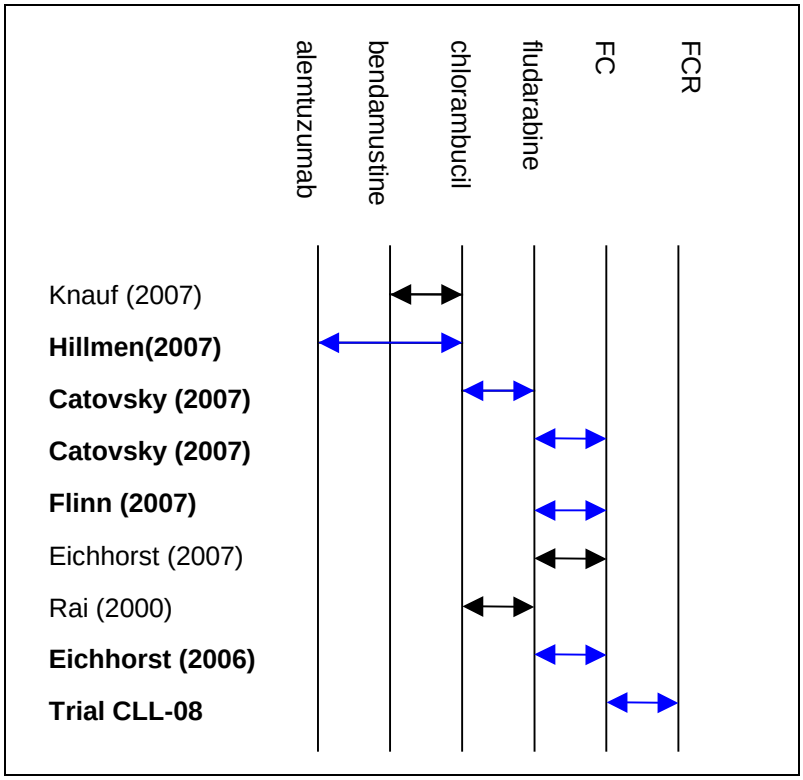
Table 13 : Summary of the trials included in the MTC for PFS

Trial	Treatment	Comparator	Hazard ratio	Lower bound	Upper bound
CLL-8[5]	R-FC	FC	0.56	0.43	0.72
Catovsky (2007)[1]	FC	Fludarabine	0.45	0.35	0.59
	Fludarabine	Chlorambucil	0.86	0.71	1.04
Hillmen (2007)[11]	Alemtuzumab	Chlorambucil	0.58	0.43	0.77
Flinn (2007)[17a]	FC	Fludarabine	0.51	NR	NR
Eichhorst (2006)[18]	FC	Fludarabine	0.56	0.40	0.80

Key: [a] Lower and upper bound not reported, stated p-value = 0.0003; Data source: Roche submission section 6.6; table 22 (p. 80)

Figure 4 provides the MTC network diagram showing the comparators included in each of the trials and how the trials are linked together to form the set of indirect comparisons underlying the model. The figure provides details of all eight trials included in the MTC for complete response and overall response, with the references in bold indicating the five trials included in the MTC for PFS.[1;5;11;17;18]

Figure 4 : Network of the five included trials in the MTC for PFS



Data source: Roche responses to ERG questions (received 18.12.08)

As chlorambucil was, on average, the treatment with the shortest PFS, this was used as the reference comparator. Table 14 shows the estimated HR of the different comparators in comparison to chlorambucil. R-FC shows the lowest HR in comparison with chlorambucil of 0.24 (lower bound: 0.17; upper bound: 0.34) implying that R-FC prolongs PFS to the greatest extent.

Table 14 : Hazard ratios with respect to chlorambucil

Treatment with chlorambucil	Mean HR	Median HR	Lower bound	Upper bound
R-FC	0.24	0.24	0.17	0.34
Fludarabine	0.86	0.86	0.71	1.04
Alemtuzumab	0.59	0.58	0.43	0.78
FC	0.43	0.43	0.33	0.55

Data source: Roche submission section 6.6; table 22 (p. 80)

Table 15 shows the estimated HR for R-FC relative to the comparators of FC alone, chlorambucil, fludarabine, and alemtuzumab. As is depicted, the upper bounds of the credibility intervals for the HR are all below 1, implying that R-FC prolonged PFS in comparison to all of these other treatments. The probability for R-FC being the best treatment of the alternatives in terms of PFS is 100%.

Table 15 : Relative efficacy of R-FC measured as HR for PFS against comparator treatments

R-FC versus treatment	Mean HR	Median HR	Lower bound	Upper bound
Chlorambucil	0.24	0.24	0.17	0.34
Fludarabine	0.28	0.28	0.20	0.38
Alemtuzumab	0.42	0.41	0.26	0.66
FC	0.56	0.56	0.43	0.72

Data source: Roche submission section 6.6; table 23 (p. 80)

Complete response

All eight trials reported the outcome of CR.[1;5;11-14;17;18] This outcome was reported in terms of the odds ratio (OR) or relative risk (RR). The results of the relative effects of each of the comparators for the outcome of CR is shown in Table 16.

Table 16 : Observed OR and RR for CR

	Treatment	OR	RR
CLL8[5]	R-FC versus FC	2.7 (2.0-3.8)	2.09 (1.63-2.69)
Catovsky (2007)[1]	FC versus fludarabine	2.3 (1.3-4.1)	2.5 (1.7-3.8)
Catovsky (2007)[1]	Fludarabine versus chlorambucil	3.5 (2.1-5.8)	2.1 (1.3-3.5)
Hillmen (2007)[11]	Alemtuzumab versus chlorambucil	15.4 (4.6-51.3)	11.9 (3.8-37.9)
Flinn (2007))[17]	FC versus Fludarabine	6.4 (2.6-15.9)	5.1 (2.2-11.9)
Rai (2000)[12]	Fludarabine versus chlorambucil	5.4 (2.4-12.1)	4.5 (2.2-9.5)
Eichhorst (2006)[18]	FC versus Fludarabine	4.3 (2.1-8.8)	3.5 (1.9-6.7)

Knauf (2007)[13]	Bendamustine versus chlorambucil	17.6 (5.3-58.5)	12.6 (4.0-39.6)
Eichhorst (2007)[14]	Fludarabine versus chlorambucil	7.2 (0.8-60.9)	6.7 (0.8-54.4)

Data source: Roche submission section 6.6; table 24 (p. 81)

The OR of 2.71 (>1) for R-FC versus FC implies that R-FC increases the complete remission percentage with respect to FC and is consistent with the results that can be drawn from the RR estimates (>1).[5] These results also indicate that chlorambucil is again the least preferable treatment, and is therefore again used as the reference treatment in the MTC. Table 17 shows the results for the OR and RR for different treatments compared with chlorambucil. From this it can be seen that R-FC shows the largest OR (mean: 31.6; lower bound: 17.5; upper bound: 53.4) and relative risk (mean 16.1; upper bound: 10.8; lower bound: 23.3) for CR compared with chlorambucil.

Table 17 : Complete response in comparison to chlorambucil

	OR				**RR**			
Treatment versus chlorambucil	Mean	Median	Lower bound	Upper bound	Mean	Median	Lower bound	Upper bound
R-FC	31.6	30.3	17.5	53.4	16.1	15.7	10.8	23.3
Bendamustine	26.2	19.7	6.5	84.6	13.6	12.4	5.4	28.9
Fludarabine	3.1	3.0	2.0	4.6	2.9	2.8	1.9	4.2
Alemtuzumab	23.2	17.2	5.8	74.8	12.6	11.4	4.9	27.7
FC	11.5	11.1	7.2	17.7	8.6	8.4	5.8	12.4

Data source: Roche submission section 6.6; table 25 (p. 81)

The OR and RR results for the relative effect on the percentage of patients in CR for each of the comparators is presented in Table 18 below. These indicate R-FC produces more CR than chlorambucil, fludarabine and FC alone. However, although alemtuzumab and bendamustine perform slightly less well than R-FC, the results are not statistically significant, and it is therefore difficult to draw conclusions regarding the relative effectiveness of the three comparators in terms of the percentage of patients in CR.

Table 18 : Relative effect on percentage of patients in complete response

R-FC versus other comparators	OR				RR			
	Mean	Median	Lower bound	Upper bound	Mean	Median	Lower bound	Upper bound
Chlorambucil	31.6	30.3	17.5	53.4	16.1	15.7	10.8	23.3
Bendamustine	1.9	1.5	0.3	5.4	1.4	1.3	0.6	2.8
Fludarabine	10.4	10.1	6.3	16.4	5.7	5.6	4.0	7.8
Alemtuzumab	2.1	1.8	0.4	6.0	1.5	1.4	0.6	3.1
FC	2.8	2.7	2.0	3.8	1.9	1.9	1.5	2.3

Data source: Roche submission section 6.6; table 26 (p. 81)

Overall response

In terms of OR, again chlorambucil was the comparator with the lowest overall response rate. The ORs for R-FC in comparison to the relevant comparators were between 2.4 (FC) and 14.8 (chlorambucil), whilst the RR for an overall response with R-FC in respect to the other comparators varied between 1.1 (FC and Alemtuzumab) and 1.8 (chlorambucil). The relevant figures showing the relative effect of R-FC in relation to all the other comparators for overall response are shown in Table 19.

Table 19 : Relative effect of R-FC regarding overall response

R-FC versus other comparators	Mean	Median	Lower bound	Upper bound	Mean	Median	Lower bound	Upper bound
Chlorambucil	14.8	14.2	8.3	24.7	1.8	1.8	1.6	1.9
Bendamustine	4.5	4.2	2.0	8.9	1.2	1.2	1.1	1.3
Fludarabine	6.0	5.8	3.5	9.8	1.3	1.3	1.2	1.4
Alemtuzumab	3.8	3.5	1.6	7.5	1.1	1.1	1.0	1.3
FC	2.4	2.3	1.6	3.4	1.1	1.1	1.0	1.1

Data source: Roche submission section 6.6; table 28 (p. 82)

4.3.2. Critique of submitted evidence syntheses

No further evidence syntheses, other than for the MTC (see Section 4.1.5.2) were presented in the clinical effectiveness section of the submission.

4.3.3. Summary of clinical effectiveness

- The submission contains all the relevant studies and the relevant data within those studies. The submitted evidence also adequately reflects the decision problem defined in the submission.
- The submission from Roche included one good quality unpublished phase III trial; the German CCL-8 study.[5]
- For the primary outcome of PFS, all patients who showed an insufficient response (i.e. stable or progressive disease) were censored at this time, and eligible to receive an alternative 2nd line treatment regimen. Patients were not necessarily crossed-over to the alternative treatment regimen in the other trial arm. Assessment of any potential differences between R-FC and FC in OS is therefore potentially confounded by the effects of 2nd-line treatment.
- There was a statistically significant increase in PFS with R-FC [median 39.8 months versus 32.2 months; HR 0.56 (95% CI: 0.43; 0.72)] compared with FC

alone. The initial treatment benefit for the R-FC regimen in terms of OS noted at the time of interim analysis (median follow-up time 20.7 months) was no longer maintained at slightly longer follow-up (median 25.4 months) [HR 0.72 (95% CI: 0.48; 1.09)]. Patients in the R-FC arm remained event free (disease progression, relapse, death or start of new CLL treatment) significantly longer than those in the FC-arm [39.8 months versus 31.1 months; HR 0.55 (95% CI: 0.43; 0.70). Response rates also significantly favoured treatment with R-FC, with 36.0% of patients in this arm achieving complete response, compared with 17.2% with FC. Partial response rates were not significantly different between trial arms at 50.1% for R-FC and 55.5% for FC respectively.

- Grade 3 or 4 adverse events were higher in the R-FC arm (77%) compared to the FC arm (62%); mostly due to a higher incidence of blood and lymphatic system disorders (57% versus 41%). Dose modifications were also more frequent in this arm. However, this did not lead to differences in treatment discontinuation. There were also no differences between arms in the rate of deaths considered related to therapy (2%).

Mixed treatment comparison

- On the basis of evidence from five RCTs, it appears that adding rituximab to FC alone significantly increased PFS compared to chlorambucil [(mean HR 0.24 (lower bound: 0.17; upper bound 0.34)], fludarabine [mean HR 0.28 (lower bound: 0.20; upper bound 0.38)], alemtuzumab [mean HR 0.42 (lower bound: 0.26; upper bound 0.66)], and FC [mean HR 0.56 (lower bound: 0.43; upper bound 0.72)].

- Addition of R to FC also significantly increased complete response compared to chlorambucil, fludarabine and FC. However, R-FC did not perform significantly better than either alemtuzumab or bendamustine in terms of complete response.

5. Economic evaluation

This chapter provides an assessment of the cost effectiveness analysis submitted by Roche. It starts with a review of the search strategy adopted by Roche and an overview of their economic model. There follows an analysis of the cost-effectiveness analysis using standard approaches for critical appraisal of economic evaluation (Section 5.2) and finally the baseline results are analysed (Section 5.3).

5.1. Overview of manufacturer's economic evaluation

5.1.1. Description of manufacturers search strategy and comment on whether the search strategy was appropriate.

Searches were performed in the following databases from 01 January 2000 to 10 November, 2008:

- MEDLINE
- MEDLINE In-Process
- EMBASE
- NHS EED
- ISPOR Research Digest

Separate search strategies were provided for each of the databases in the manufacturer's submission. Unlike the clinical searches, the economic searches included text and thesaurus terms in MEDLINE and EMBASE for the population and intervention. A thesaurus based study design filter has also been added to the MEDLINE and EMBASE searches which while not as inclusive as one with both thesaurus and text words is appropriate. There are no additional limits or filters on any of the search strategies. The ISPOR and NHS EED searches use text words combining rituximab and CLL. The combination of population and intervention is appropriate but the full spelling out of the population to "chronic lymphocytic leukaemia" would have been expected.

5.1.2. Model Structure

The modelling approach adopted by Roche uses a Markov state-transition cost-utility model implemented in Microsoft Excel. A three state model is used with a cycle length of one month and a 'lifetime' time horizon (equating to 15 years after which less than 1.5% of patients survive). The three states deployed in the model are; Progression Free survival (PFS), Progressed, and Death. All patients are assumed to start in the PFS state (defined with reference to the CLL-8 study). After each cycle a patient in PFS will experience one of three outcomes; 1) remain in the PFS state, 2) move to the Progressed state, or 3) die. Once within the Progressed state patients will either remain in this state or die. Death is an absorbing state. There is no provision within the model for patients to move back to the PFS state after moving to the Progressed state. The Progressed state therefore acts to aggregate all events subsequent to first treatment relapse (except death).

5.1.3. Natural history

The model submitted by Roche follows an established and simple structure for modelling cancer treatments which separates the progression of the illness into the three states:- Progressive Free State (PFS), Progressed State, and Death. As stated, a key simplifying assumption implicit in the Roche modelling structure is the absence of any transition from the progressed state back to PFS. This seems specifically important in the context of first line treatment of CLL since the model assumes the aggregation of all outcomes after relapse (apart from death) into a single progressed state.

5.1.4. Treatment effectiveness within submission

Treatment effectiveness in a state transition modelling framework is primarily represented by the transition probabilities between states. For these, the Roche submission relies almost entirely on outcome data from the CLL-8 trial. The only exception is the use of Office of National Statistics data to inform the background mortality rate from the PFS state for patients which is used in conjunction with the rate derived from the CLL-8 study (as described below).

For the R-FC v FC comparison, the probability of patients remaining in the PFS state in each arm of the model is time-dependant and, in the base case, derived from a Weibull parameterisation of the PFS survival rates taken from the respective arms of the CLL-8 trial. The probability of death within the PFS state is calculated as the maximum value of two sources; 1) the monthly probability experienced by patients in the CLL-8 trial (0.001196 for R-FC and 0.001388 for FC) 2) the age specific background mortality rate taken from UK life tables.

The comparison of R-FC v. chlorambucil deploys the multiple treatment comparison (reviewed in Section 4.1.5.2 above) to derive an estimate for the hazard ratio of 0.24 between the arms in the model for transition from PFS to Progressed state. Mortality rates, both from the PFS state and the Progressed state, in this comparison use the equivalent values to the RF-C v. FC comparison.

For both R-FC v. FC and R-FC v. chlorambucil, the assumption is made of equivalent adverse event profiles across arms. No incremental effects or adverse events are modelled in the base case, although alternative assumptions about adverse events are explored in sensitivity analyses.

5.1.5. Health related quality of life

The model submitted by Roche requires three utility parameters to represent quality of life in each of the three states of the model. Death, by convention, is assumed to be zero. The remaining two utility values are for the PFS and Progressed states. For these states, utility values of 0.8 and 0.6 respectively have been adopted. These are based on those generated for a HTA report by Hancock and colleagues on the use of fludarabine as first line treatment for CLL[32] and were estimated by the report authors. They based their estimates for the PFS and Progressed health states in first line CLL on the mean global quality of life scores of 81 CLL patients collected using the EORTC-QLQ-C30)[30] and Functional Assessment of Cancer – General (FACT-G) in a study by Holzner and colleagues (2001).[33] The utility values used in the model are therefore not preference-based, but represent non-clinical author estimates based crudely on condition-specific data. In their submission, Roche address this departure from the NICE reference case and cite a separate study which they have commissioned. No results from this study have been made available to the ERG.

5.1.6. Resources and costs

The perspective adopted in the Roche analysis is the UK NHS and Personal Social Service as specified in the NICE reference case.

5.1.7. Discounting

A discount rate of 3.5% has been applied to both costs and benefits within the model in accordance with the specification of the NICE reference case.

5.1.8. Sensitivity analysis

Both one-way (univariate) sensitivity analysis and probabilistic sensitivity analyses are reported in the Roche submission. Univariate sensitivity analysis is conducted on the following data parameters:

- variations in the PSA survival function
- dose levels of FC
- utility difference between PSA and progressed states
- adverse event costs
- supportive care costs
- drug administration costs.

In addition to these univariate analyses, Roche include two separate scenario analyses in their submission. These explore the effects firstly, of intravenously rather than orally administered FC, and secondly, the application of the model for cost-effectiveness comparisons of rituximab in combination with other chemotherapies.

5.1.9. Model validation

5.1.9.1. Internal Validation

No computational errors were detected when the model was tested for logical consistency, and when checking the mathematical methods used for curve fitting. Roche also report that they commissioned an independent internal validation of the model according to the criteria of: completeness of reported results and extreme testing to check plausibility of model outcomes.

5.1.9.2. External Validation

Roche report external validation of their model with reference to a comparison with the MD Anderson study reported in Tam et al 2008.[7] This study, in common with CLL-8, compared patients treated with R-FC v. FC. The Roche model was modified based on the data from this study, the main difference being the post-progression probability of death, which was four times lower in the MD Anderson study than that observed in the CLL-8 trial. Roche report that the incremental life years and QALYs predicted by their economic model is broadly consistent when populated with either of these data sets (i.e. MD Anderson study or CLL-8). Summary results from these two scenarios are presented in Table 20 below.

Table 20 : Comparison of outcome measure in the economic model based on the MD Anderson Study and the UK economic model based on CLL-8 trial data

	MDA results	UK CLL-8 results
Total FC life years	7.56	4.65
PFS life years	4.32	2.93
Progression life years	3.24	1.73
Total R-FC life years	8.89	5.73
PFS life years	6.36	4.11
Progression life years	2.54	1.62
Incremental life years	**1.34**	**1.07**
Total FC QALYs	5.4	3.38
PFS QALYs	3.45	2.34
Progression QALYs	1.95	1.04
Total R-FC QALYs	6.61	4.26
PFS QALYs	5.08	3.29
Progression QALYs	1.52	0.97
Incremental QALYS	**1.21**	**0.88**

Data source: Roche submission section 7.2.13; table 55 (p. 142)

5.2. Critique of approach used

In this section, we critically appraise the cost-effectiveness analysis submitted by Roche. First, we consider the economic model against checklists of good practice, then we critically appraise the model structure and data inputs.

5.2.1. Critical appraisal frameworks

We considered the economic evaluation submitted by Roche within the following frameworks: NICE Reference Case[34] (Table 21), Drummond and colleagues[26] (Table 22) and Philips and colleagues[35] (Table 23).

Table 21 : Critical appraisal checklist based on NICE Reference Case requirements

NICE reference case requirement		Critical Appraisal	Reviewer comment
Defining the decision problem	The scope developed by the Institute	✓	
Comparator	Therapies routinely used in the NHS, including technologies regarded as current best practice	✓	The model compares the new intervention against the two most commonly used conventional treatments (FC and chlorambucil).
Perspective on costs	NHS and PPS	✓	
Perspective on outcomes		✓	
Type of economic evaluation	Cost-utility analysis	✓	
Synthesis of evidence on outcomes	Review of evidence is given.	✓	Mainly based on evidence of single relevant trial CLL-8 (MTC uses other evidence to derive values for the R-FC v chlorambucil comparison)
Measure of health benefit	QALYs	✓	
Source of data for measurement of HRQL	Uses values based on author estimates only.	✘	Further dedicated study is mentioned as on-going and due to report in Jan 09 but results have not been provided to the ERG.
Source of preference data for valuation of changes in HRQL		✘	Not empirically based
Discount rate	3.5% per annum for costs and benefits	✓	
Equity weighting	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit.	✓	

✓ indicates 'Clear', ✘ indicates 'concerns', ? indicates 'unclear/unknown'.

Table 22 : Critical appraisal checklist from Drummond and colleagues

Item	Critical Appraisal	Reviewer comment
Is there a well defined question?	✓	
Is there a clear description of alternatives (i.e. who did what to whom, where, and how often)?	✓	R-FC v. FC R-FC v chlorambucil
Has the correct patient group/population of interest been clearly stated?	✓	The model uses a cohort with patient characteristics based on the CLL-8 study
Is the correct comparator used?	✓	Major comparators are considered in the analysis. Some comparators are not considered (eg. MabCampath).
Is the study type reasonable?	✓	
Is the perspective of the analysis clearly stated?	✓	The analysis takes the perspective of the NHS and Personal Social Services
Is the perspective employed appropriate?	✓	
Is effectiveness of the intervention established?	✓	CLL-8 study shows a clear advantage in prolonging Progression Free Survival although its effectiveness in prolonging overall survival is less clear.
Has a lifetime horizon been used for analysis, if not has a shorter time horizon been justified?	✓	Model time horizon is 15 years after which only about 1.3% of the population cohort and modelled as surviving.
Are the costs and consequences consistent with the perspective employed?	✓	Costs are presented from the NHS and PSS perspective.
Is differential timing considered?	✓	
Is incremental analysis performed?	✓	
Is sensitivity analysis undertaken and presented clearly?	✓	Univariate and probabilistic sensitivity analysis are presented.

✓ indicates 'Clear', ✘ indicates 'concerns', ? indicates 'unclear/unknown'.

Table 23 : Critical appraisal checklist against Philips and colleagues framework for decision analytic modelling practice

Dimension of quality			Comments
Structure			
S1	Statement of decision problem/objective	✓	
S2	Statement of scope/perspective	✓	
S3	Rationale for structure	✘	Reference to established approaches to modelling cancer are used to justify approach. However, implications of structural uncertainty

			are not fully explored.
S4	Structural assumptions	✘	A key limitation is the simplifying structural assumption of an aggregated progressed state in the model. These assumptions should be explored more thoroughly.
S5	Strategies / comparators	✓	The two most prevalent treatments (FC and chlorambucil) are modelled as comparators.
S6	Model Type	✓	A conventional state-transition Markov model was used.
S7	Time horizon	✓	The 'lifetime' 15 time horizon was long enough. In addition the model included the facility to change the time horizon.
S8	Disease states / pathways	?	The aggregation of post-relapse into one progressed state may be an over-simplification especially given this is for first-line treatment
S9	Cycle length	✓	A cycle length of one month was employed which is appropriate for this analysis.
Data			
D1	Data identification	✓	
D2	Pre-model data analysis	✓	
D2a	Baseline data	✓	
D2b	Treatment effects	✓	MTC used to derive treatment effect for the chlorambucil comparison
D2c	Quality of life weights (utilities)	✘	Chosen values not based on empirical study.
D3	Data incorporation	✓	
D4	Assessment of uncertainty	✘	Not all aspects of structural uncertainty are addressed in the analysis.
D4a	Methodological	✓	
D4b	Structural	?	Alternative structures might be considered given that this is first-line treatment.
D4c	Heterogeneity	✓	Standardised cohort based on
D4d	Parameter	✓	Probabilistic and univariate sensitivity analysis have been performed
Consistency			
C1	Internal consistency	✓	Tests of the mathematical consistency of the model showed no serious errors.
C2	External consistency	✓	Evidence is presented of both parameter validation and model validation.

✓ indicates 'Clear', ✘ indicates 'concerns', ? indicates 'unclear/unknown'.

5.2.2. Modelling approach and structure

The following sections examine the approach adopted by Roche to develop and parameterise the economic model.

5.2.3. Data Inputs

5.2.3.1. Patient Group

The modelled patient cohort is based on the trial population used in the CLL-8 trial (described in Section 6. p40-46 in the Roche submission). The mean age of the population is 59 years and the model assumes a male to female ratio of 1.6.[36] Use of this patient group is justified as matching the cohort used in the CLL-8. No sub-group analysis is modelled in the analysis.

5.2.3.2. Clinical Effectiveness Data

Clinical effectiveness parameters in the model are incorporated through the transition probabilities between states. These values and the methods used to derive them are reviewed below.

5.2.3.2.1. R-FC v. FC comparison

For the R-FC v. FC comparison, data values are almost exclusively drawn from the CLL-8 clinical trial which forms the primary basis for the submission.

Progression Free Survival

The most prominent effectiveness parameter in the model is Progression Free Survival between the arms of the model. For this, the R-FC v. FC base case adopts a method of curve fitting using a Weibull parameterisation to extrapolate from the CLL-8 clinical data (p.60-61 of the Roche submission). Weibull parameters were estimated using patient level data from the study, and since stratified and unstratified outputs were consistent, an unstratified model was used to determine the Weibull parameters for each arm of the comparison.

Given the importance of this parameter in the model, Roche assess a range of different curve parameterisations for PFS survival curve for goodness of fit. They demonstrate a

clear rationale for the adoption of a Weibull fit in preference to several alternative parametric survival functions: Exponential, Log Logistic, Log Normal and Gompertz (p. 115-117 of the submission). In general, we are satisfied that Roche have used an appropriate method both (a) to extrapolate progression-free survival from the clinical data and (b) to determine the best curve parameterisation.

PFS-Death

To model the probability of death for patients in the PFS state, the model uses the *maximum* of the two rates; 1) the monthly rate derived from the CLL-8 trial for each of the arms (0.001196 for R-FC and 0.001388 for FC) and 2) the age specific background mortality rate taken from UK life tables and adjusted for the gender ratio used in the model. The second rate is used to ensure that the mortality rate for patients in this state in the model does not fall below that of the average UK population which seems clinical implausible. The very small difference in monthly mortality rate from PFS between the two arms found in the CLL-8 contributes only a negligible benefit advantage to R-FC in the model.

Progressed to Death

Patients in the Progressed state were modelled as a single population, with no distinction in the probability of death between arms of the model. The rationale given for this is that there was no significant difference found between arms in the CLL-8 trial. For this probability a constant value was calculated based on the inverse of the mean from the CLL-8 trial Kaplan-Meier. This resulted in a monthly probability of dying in this state of 0.0405. This probability of death from the Progressed state is used throughout the time horizon of the model.

5.2.3.2.2. R-FC v. Chlorambucil

For the R-FC v. chlorambucil comparison the MTC analysis (as described in section 4.1.5.2 above) is used in the model to derive a hazard ratio of 0.24 between the arms. This result is validated against data from CLL4 study[1]. Calculation of death probabilities from the PFS and progressed state in this comparison use the approach described above for the R-FC v. FC comparison.

5.2.3.3. Drug costs

Costs for the drugs used in the model are consistently drawn from the British National Formulary (BNF 56) and we are satisfied that these values are correct and have been properly applied in the model.

Calculation and explanation of the dose levels and associated costs of treatment are provided in the report and we are satisfied that this represents a valid approach.

The drug doses and costs for rituximab are outlined in Table 24, and, for FC and chlorambucil, in Table 25 below.

Table 24 : Drug doses and costs for rituximab

Assumptions	Value	Description
Body surface area m2	1.93	Average body surface area (from CLL-8)
Unit price per vial (£) 100mg 500mg	 174.63 874.15	BNF 56
Recommended dose (mg/m2) Cycle 1 Cycle 2-6	 375 500	Recommended adult dose as per SPC
Average adult Dose (mg) including wastage Cycle 1 Cycle 2-6	 800 1,000	 1.93mg/m2 * 375mg = 725mg (round-up) 1.93mg/m2 * 500mg = 965mg (round-up)
Cost per infusion/cycle (£) Cycle 1 Cycle 2-6	 1,397.03 1,746.30	 500ml @ £873.15 + 3* 100ml @ £174.63 2 * 500ml @ £873.15
Number of infusions cycles	6	Administered on day 0 in Cycle 1 and day 1 of each subsequent cycle of chemotherapy in 28 day cycles for a total of 6 cycles
Total rituximab drug cost per patient (£)	10,128.53	£1,397.04 + 5 * £1,746.30

Data source: Roche submission section 7.2.9.2; table 45 (p. 128)

Table 25 : Drug doses and costs for FC and chlorambucil

Assumptions	F (oral)	C (oral)	Chl	Description
Body surface area m2	1.93	1.93	1.93	Average adult body surface area
Unit price per mg (£)	1.86	0.0024	0.17	BNF 56
Recommended dose	24	150	10	Recommended adult dose

(mg/m2)				
Average adult daily dose (mg) including wastage	50	300	20	F: 24mg*1.93m2 = 46.32mg (round to nearest 10mg) C: 150mg*1.93m2 = 289.50mg (round to nearest 50mg) Chl: 10mg*1. 93m2 = 19.3mg (round to nearest 2mg)
Days of treatment per cycle	5	5	7	Recommended adult dose
Cost per infusion/cycle	465.00	3.60	23.80	F: £1.86 * 50mg/day *5 days C: £0.0024 * 300mg/day *5 days Chl: £0.17 * 20mg/day *7 days
Number of cycles of treatment	6	6	12	Administered on day 1 of each cycle of chemotherapy in 28 day cycles
Total drug cost per patient (£)	2,790	21.60	285.60	F: £465 * 6 cycles C: £3.60 * 6 cycles Chl: £23.80 * 12 cycles

Data source: Roche submission section 7.2.9.2; table 45 (p. 128)

5.2.3.4. Disease management costs

The model takes an NHS and PSS perspective, which is appropriate. Values for costs associated with administration of R-FC and the comparators are taken from the appropriate reference costs for inpatient chemotherapy administration in National Schedule of Reference Costs 2006-07. These are shown in Table 26 below.

Table 26 : Summary of Drug administration costs used in the model

Applied to:	**HRG label (Code)**	**National average unit costs**
FC (oral) on day 1 of each cycle	Deliver exclusively Oral Chemotherapy (SB11Z)	£280
R (in combination with FC or on its own during cycle 1)	Deliver complex Chemotherapy, including prolonged infusional treatment at first attendance (SB14Z)	£430
Sensitivity analysis: FC infusion for day 1	Deliver simple Parenteral Chemotherapy at first attendance (SB12Z)	£309
Sensitivity analysis: FC infusion for days 2-3	Deliver subsequent elements of a chemotherapy cycle (SB15Z)	£255

Data source: Roche submission section 7.2.9.2; table 46 (p. 128)

The base case assumes that rituximab incurs a marginal cost of £430 on day 0 of cycle 1 in the model. The patient is then assumed to return to collect their oral FC therapy (or chlorambucil) for a cost of £280. In the subsequent 5 cycles in the R-FC arm of the model patients are assumed to enter the hospital on day 1 and receive the rituximab infusion and collect the FC (or chlorambucil) treatment for subsequent therapy. This leads to a marginal cost for rituximab of £430 in the first cycle and £150 for cycles 2-6 in the model. In addition, a scenario analysis is presented (see Section 5.3.1.3 below) which examines the effect of using intravenous administration of FC (as used in the CLL-8 trial).

In addition to the drug administration costs, a face-to-face consultation cost of £84 is included into the model at each cycle. This is the same for both arms.

Subsequent 2nd line therapies for CLL treatment were used to derive an average cost for post-progression of £257.66 per monthly cycle which was applied to the Progressed state in the model. This value was derived by aggregating unit costs from BNF 56 (standard doses) for those therapies representing resources used by more than 2% of the progressed patient population based on the collected CLL-8 trial data. Table 27 shows those therapies which were used to determine these costs. Because progressed patients were treated as a single population in the model the average cost (£257.66 per monthly cycle) was applied equally across arms. No sensitivity analysis was provided to explore the effects of *differential* costs between model arms in the Progressed state.

Table 27 : Subsequent CLL treatments from the CLL-8 trial used to derive costs for the modelled progressed state

Progression Therapy	Number of Patients		Cumulative days on treatment		Average number of days per patient on treatment	
	FC	R-FC	FC	R-FC	FC	R-FC
ALEMTUZUMAB	8	5	550	251	68.75	50.2
ALEMTUZUMAB/CYCLOPHOSPHAMIDE/FLUDARABINE	1	2	120	145	120	72.5
CHLORAMBUCIL	1	3	5	157	5	52.33
CYCLOPHOSPHAMIDE	3	0	33	0	11	0

CYCLOPHOSPHAMIDE/DOXORUBICIN/PREDNISOLONE/RITUXIMAB/VINCRISTINE	12	7	881	213	73.42	30.43
CYCLOPHOSPHAMIDE/DOXORUBICIN/PREDNISONE/VINCRISTINE	9	1	645	105	71.67	105
CYCLOPHOSPHAMIDE/FLUDARABINE/RITUXIMAB	4	2	287	146	71.75	73
FLUDARABINE	2	1	32	195	16	195
RADIOTHERAPY	2	0	18	0	9	0
RITUXIMAB	5	2	155	84	31	42
STEM CELL TRANSPLANT	1	4	9	0	9	0

Data source: Roche submission section 7.2.9.2; table 48 (p. 131)

5.2.3.5. Adverse event costs

The model base case assumes that the adverse event profile is the same for both arms in both R-FC vs FC and R-FC v. chlorambucil comparisons. This assumption is justified by reference 1) to the trial evidence from the CLL-8 study which shows no statistically significant differences between arms in the incidence of adverse events between R-FC v FC, and (2) by the absence of comparable data for R-FC and chlorambucil. A range of severe adverse events are included in the sensitivity analysis including febrile neutropenia, blood transfusions and bone marrow transplants. This sensitivity analysis, however, applies the adverse event profile equally across arms and it does not explore *differential* rates of adverse events across arms despite the fact that (statistically non-significant) differences were found in the incidence of severe adverse events in CLL-8 (17 versus 24 Grade 3, and 8 verses 15 Grade 4 events for FC v. R-FC respectively).

5.2.3.6. Health related quality of life

As detailed in Section 5.1.4, two health state utilities (apart from death which has a utility of zero) were required in the model. These values were based on a previous study by Hancock and colleagues[37] and were derived from author estimates. In addition, Roche reference a further on-going quality of life study that has been commissioned (section 7.2.8.3, p.125 of the Roche submission). However, no further information has been provided to the ERG about this study.

5.2.4. Assessment of uncertainty

Roche present a range of one-way, scenario, and probabilistic sensitivity analyses in their submission. The main reported findings from the analysis of parameter sensitivity in the Roche submission are summarised below:

- The greatest changes to the cost-effectiveness outputs are caused by changes to the parametric functions used for the extrapolation of PFS survival curves (e.g. use of Gompertz or exponential functions) and also when large changes are made to the utility differential between the states.
- ICER values output in the sensitivity analysis remained within normally accepted willingness to pay thresholds for cost-effectiveness.
- Probabilistic Analysis tends to support the findings of the deterministic model that the ICER from the model is very unlikely to exceed normally accepted cost-effectiveness thresholds.

These findings are discussed in more detail in the following section. We also outline a concern relating to the treatment of structural uncertainty in the model.

5.3. Results included in manufacturer's submission

Here we present a summary of the results from the Roche submission. See Section 7.3 p.143-173 of the Roche submission for further details.

5.3.1. Deterministic Results

5.3.1.1. Base Case

Base case results from the two modelled comparisons are shown in Table 28. These show ICERs of £13,189 and £6,422 for the R-FC v. FC and R-FC v. chlorambucil comparisons respectively.

Table 28 : Base case result from the Roche model for life year/cost per QALY gained ratios for R-FC versus FC and R-FC versus Chlorambucil

Cost-utility results : R-FC v. FC	R-FC	FC	Incremental
Mean Life Years (yrs)	5.73	4.65	1.07

Mean QALYs	4.26	3.38	0.88
Mean Total Cost	£25,595	£13,978	£11,617
Cost per Life Year Gained (£)			**£10,825**
Cost per QALY Gained (£)			**£13,189**
Cost-utility results : R-FC v. chlorambucil	**R-FC**	**Chlorambucil**	**Incremental**
Mean Life Years (yrs)	5.73	3.40	2.33
Mean QALYs	4.26	2.35	1.91
Mean Total Cost	£25,595	£13,345	£12,250
Cost per Life Year Gained (£)			**£5,253**
Cost per QALY Gained (£)			**£6,422**

Data source: Roche submission section 7.3.1.1; tables 58 and 61 (p. 144-6)

5.3.1.2. One-way sensitivity analysis

A range of one-way sensitivity analyses are presented in the Roche report which are summarised in Table 29 below. These show that the model output is most sensitive to changes in the parameterisation of the PFS survival curve and to large fluctuations in the utility levels in the model.

Table 29 : One-way sensitivity analysis summary results from the Roche model for life year/cost per QALY gained ratios for R-FC versus FC and R-FC versus chlorambucil

Category	Sensitivity analyses	R-FC v. FC ICER
Base case		**£13,189**
Curve fitting for PFS Survival function	Exponential function	£10,249
	Log logistic function	£13,164
	Log normal function	£12,426
	Gompertz function	£22,661
Dose levels of FC:	IV Infusion of FC = Actual dose from trial	£12,236
	IV Infusion of FC = Recommended dose	£13,006
Utility difference between states :	Utilities: PFS=0.9; Progressed = 0.5	£11,497
	Utilities: PFS=0.75; Progressed = 0.65	£14,236
Adverse events :	Inclusion of adverse event costs	£13,283
Supportive care costs :	Monthly supportive care cost increase by 50%	£13,155
	Monthly supportive care cost decrease by 50%	£13,223
Drug administration costs :	Drug administration cost upper quartile	£14,519
	Drug administration cost lower quartile	£12,298

Category	Sensitivity analyses	R-FC v. Chl ICER
Base case		**£6,422**
Utility difference between states :	Utilities: PFS=0.9; Progressed = 0.5	£5,612
	Utilities: PFS=0.75; Progressed = 0.65	£6,921
Adverse events :	Inclusion of adverse event costs	£6,756
Supportive care costs :	Monthly supportive care cost increase by 50%	£6,419
	Monthly supportive care cost decrease by 50%	£6,424
Drug administration costs :	Drug administration cost upper quartile	£6,400
	Drug administration cost lower quartile	£6,344

Data source: Roche submission section 7.3.3.1; table 62 and 63 (p. 146-7)

5.3.1.3. Scenario Analyses

In addition to the univariate sensitivity analyses provided in the submission, Roche conducted two scenario analyses to investigate the effects on the model relating to specific treatment configurations.

The first scenario analysis investigates the effects of FC administered intravenously rather than orally. Two changes are made to the model. Firstly, the cost of oral FC treatment is replaced with the cost of i.v. FC treatment, and secondly the recommended dose levels are replaced by those used in the CLL-8 trial. Both these changes cause a decrease in the incremental cost between the arms in the R-FC v. FC comparison, there is no impact on incremental QALYs. The resultant impact is to slightly reduce the ICER to £12,236 per QALY (from £13,189). For the R-FC v. chlorambucil comparison the impact of the changes is to increase the incremental cost of the R-FC arm but this effect is minimised because there is also a reduction in dose which cancels out much of the incremental administration costs. The overall impact is to increase the ICER to £7,424 per QALY (from £6,422).

The second scenario analysis investigates the use of the model as a basis for determining the cost-effectiveness of rituximab in combination with other appropriate chemotherapy combinations. Examples of such combinations which have been the subject of Phase II trials include rituximab used in combination with fludarabine alone, rituximab with pentostatin and cyclophosphamide, and R-FC plus mitoxantrone. A direct economic analysis is not possible in these instances due to lack of data.

However, based on an assumption of equivalent incremental costs between another chemotherapy regimen and the incremental costs modelled for the R-FC v FC comparison, a threshold analysis is presented which illustrates the effect of reducing the level of QALY gain in the model on the ICER output value. This threshold analysis shows that reduction of QALY gain to about 40% of that modelled for the R-FC v. FC analysis would be required to produce an ICER above the £30,000 per QALY threshold for willingness-to-pay. Based on this it is argued in the submission that rituximab is likely to remain cost-effective with other chemotherapy combinations.

5.3.2. Probabilistic Results

The mean values obtained in the probabilistic analysis in the model are shown in Table 30 below and match the deterministic results very closely.

Table 30 : Mean Cost Effectiveness results for R-FC versus FC and R-FC versus chlorambucil (1000 runs)

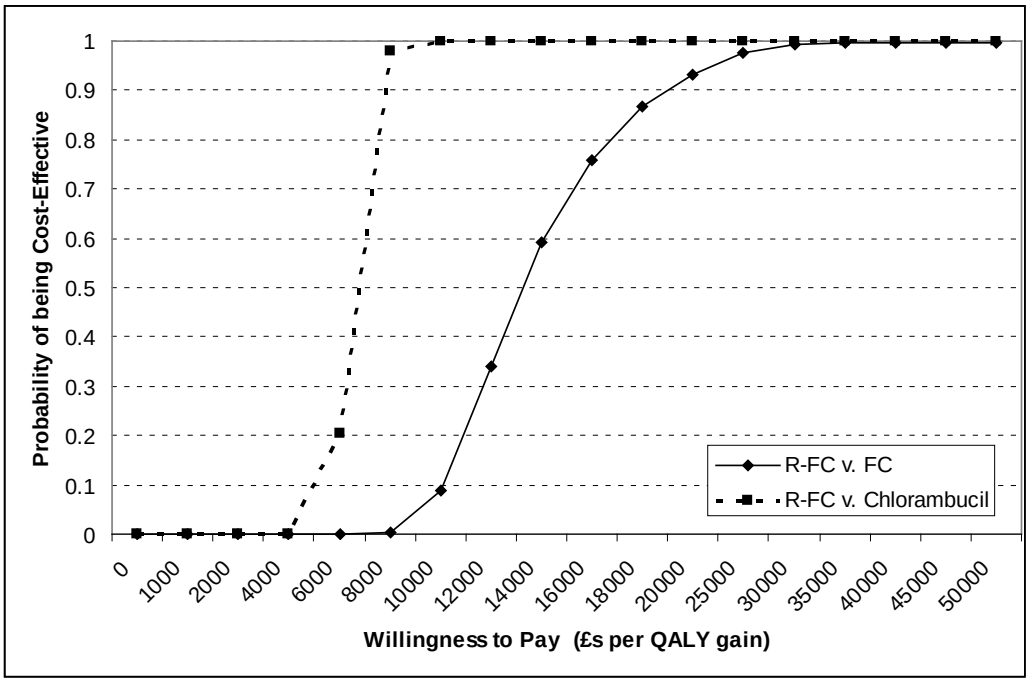
Cost-utility results (R-FC v. FC)	**R-FC**	**FC**	**Incremental**
Mean Life Years (yrs)	5.73	4.66	1.07
Mean QALYs	4.26	3.38	0.88
Mean Total Cost	£25,691	£14,002	£11,689
Cost per Life Year Gained (£)			**£10,916**
Cost per QALY Gained (£)			**£13,295**
Cost-utility results (R-FC v. Chlorambucil)	**R-FC**	**Chl**	**Incremental**
Mean Life Years (yrs)	5.72	3.41	2.31
Mean QALYs	4.26	2.37	1.89
Mean Total Cost	£25,536	£13,289	£12,247
Cost per Life Year Gained (£)			**£5,302**
Cost per QALY Gained (£)			**£6,482**

Data source: Roche submission section 7.3.3.1; tables 68 and 69 (p. 152)

The PSA presented in the report also largely confirms the impression given by the univariate sensitivity analysis and shows a strong probability that R-FC is cost-effective at both £20,000/QALY and £30,000/QALY willingness-to-pay thresholds (91.9% at £20,000 and 98.6% at £30,000 thresholds for the R-FC v. FC comparison, and 100% for both thresholds for the R-FC v. Chlorambucil comparison). The outputs from the probabilistic simulation in the model are most clearly illustrated in the cost

effectiveness acceptability curves shown in Figure 5 below which are generated from the simulation in the PSA analysis.

Figure 5 : Cost Effectiveness Acceptability curves for both R-FC v. FC and R-FC v. chlorambucil comparisons (examples based on 1000 model simulations)



Data source: Roche submission section 7.3.3.1; Figures 29 and 30 (p. 154)

5.4. Comment on validity of results presented with reference to methodology used

In general, the modelling methods, construction and parameterisation of the model are well explained in the Roche submission. In our testing of the model, we found no errors relating to logical consistency or in the mathematical methods used to calculate parametric values.

The main limitations of the economic analysis and modelling are:

- Almost all data parameters for effectiveness are drawn from a single study – CLL-8. Although this study seems of good quality, there are clearly limitations inherent in an analysis which relies on a single clinical trial.

- Utility values in the model are not based on a methodologically appropriate source.

- There is a fundamental issue of structural uncertainty in the model relating to the treatment of overall survival rates between arms which has not been explored in the analysis. This issue is discussed in more detail below.

5.5. Summary of uncertainties and issues

The economic model and analysis presented by Roche in their submission is generally well presented and provides a case for a stable ICER which falls below accepted willingness-to-pay thresholds against FC and chlorambucil in the treatment of CLL.

The univariate and probabilistic sensitivity analyses show that the base case ICER remains consistently below £20,000 per QALY and the PSA shows a very high probability that the use of R-FC is cost-effective at £20,000 and £30,000 per QALY thresholds.

However, the investigation provided by Roche concentrates on the parameter uncertainty in the model and key issues of structural uncertainty are not fully addressed. In this context, we would highlight particularly the model assumption of aggregation in the post-relapse state, which is acknowledged in the report as a limitation (Section 7.3.4.3 p.156). The aggregation is clinically unrealistic because patients will receive further treatment at progression which may then result in further periods of progression-free survival. The relapsing nature of CLL means that subsequent periods of progression are less likely to respond to further treatment, implying that later periods of progression in the course of disease are likely to be associated with higher disease-related mortality. This casts doubts over the simplifying assumption of a constant hazard of death after progression as modelled by Roche. This assumption was not confirmed through exploration of the CLL-8 dataset, as was done comprehensively for the modelling of PFS.

The overall effect of the aggregated Progressed disease state and constant hazard of death from this state is to imply a correlation between PFS and OS which we do not believe has been empirically demonstrated. This is because, while successive cycles

of chemotherapy will remove chemosensitive malignant cells, proliferation of cross-resistant clones continues, ultimately leading to further relapse and death. This accounts for the failure of fludarabine to demonstrate overall survival gains despite initially promising results on progression free survival. The potential for rituximab to address cross-resistance is currently unknown, partially reflecting incomplete understanding of its mechanism of action. We therefore believe that the implications of this important structural assumption on model outputs should be explored in sensitivity analysis.

In addition, it should be borne in mind that once any assumed benefit in overall survival is removed from the model then the benefit outputs become highly sensitive to the utility parameters assumed for the PFS and Progressed states and these utility values are not currently available from an appropriate source.

Following queries on this subject by the ERG, Roche have carried out an analysis which differentially increases the probability of death for the R-FC arm for the Progressed state of the model. The outcome of this analysis shows the effect of removing the impact of OS advantage in the R-FC arm of the model. At the limit of this analysis (where no OS advantage is generated by the model structure) the ICER rises to £30,336. At this limit it should be noted that the ICER becomes very sensitive to the level of utilities used for the PFS and Progressed states of the model.
Given our concerns about this aspect of structural uncertainty within the model we have included some additional analysis in the next section.

6. Additional work undertaken by the ERG

6.1. Clinical effectiveness

No further work was undertaken by the ERG for the clinical effectiveness sections of this report.

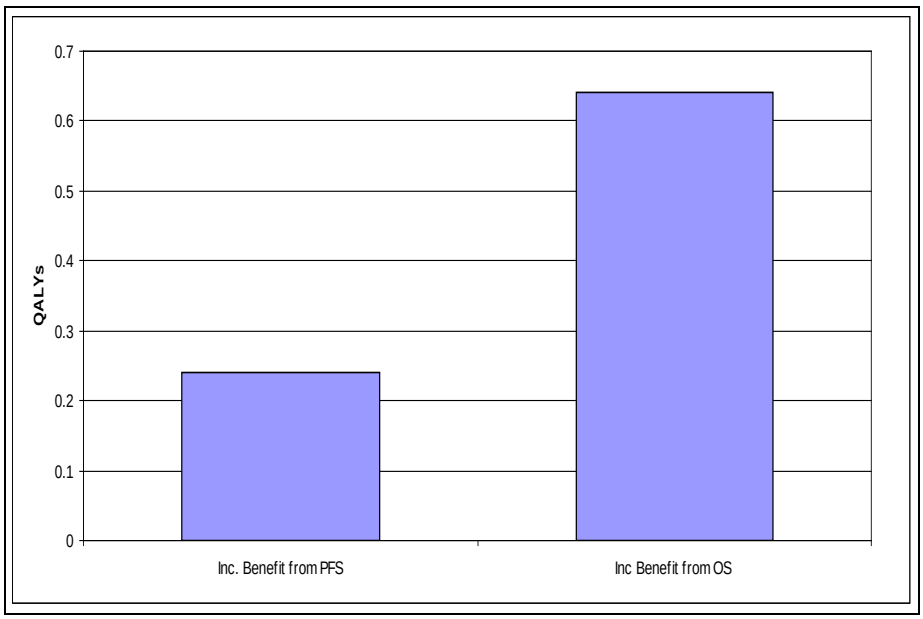
6.2. Cost Effectiveness Analysis

In this section we present some additional analyses related to the economic modelling and cost-effectiveness. We have focussed here on R-FC v. FC since this is the comparison of primary interest and the one which is based directly on the CLL-8 trial data.

6.2.1. Component Analysis of Utility gains

The two major sources of incremental utility generated by the model correspond to benefit that can be demonstrated 1) by an advantage in PFS, and 2) an advantage in OS driven by earlier exposure in the FC arm to the risk of mortality from progressive disease. To assess the relative contribution of these sources a simple component analysis was conducted. The utility advantage of the PFS versus the Progressed state in the model was removed (i.e. both states were given a utility of 0.6). The remaining incremental QALYs between arms was then assumed to derive solely from the OS benefit and the difference between this value and the Roche base case assumed to be due to the utility difference between the PFS and Progressed states. Figure 6 shows the relative contribution, based on this method, of these two sources of benefit gain in the base case model for the R-FC v. FC comparison (0.24 QALYs for PFS and 0.64 QALYs for OS).

Figure 6 : Component Analysis showing the relative contribution to incremental utility from different factors in the model for R-FC v. FC. Deterministic base case outputs.

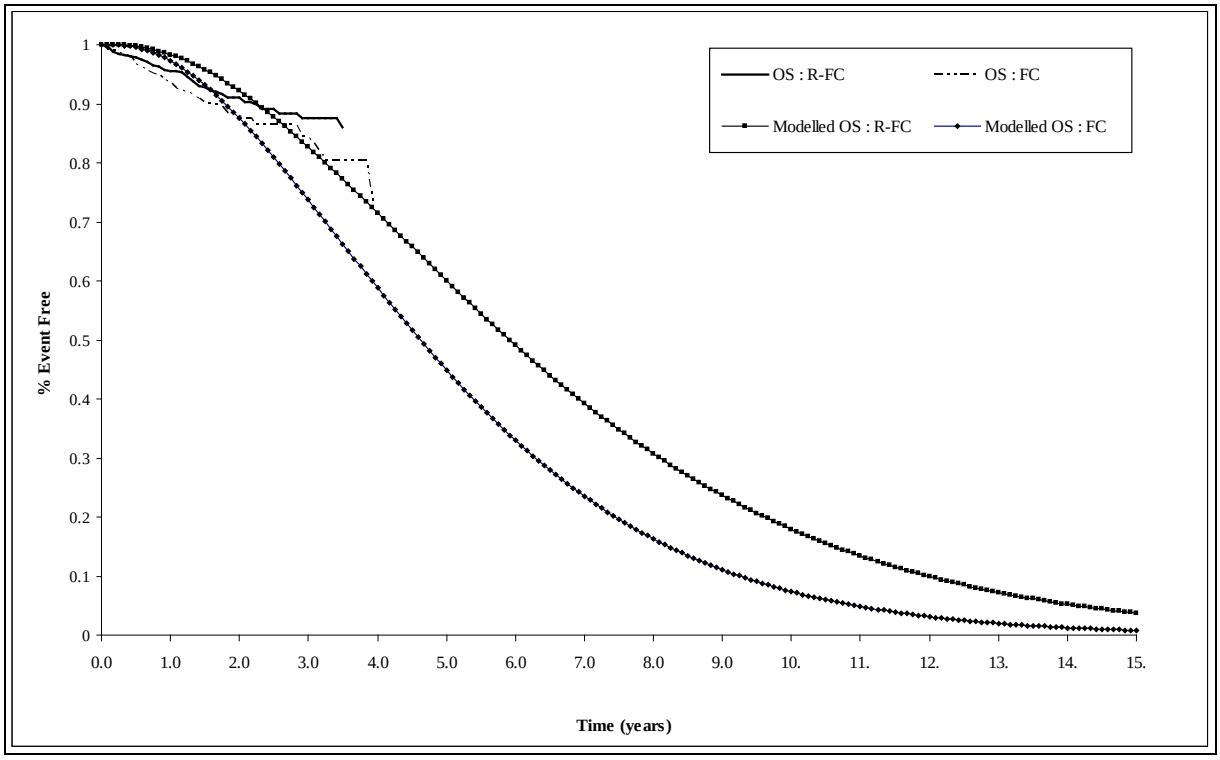


From this analysis it is clear that majority of incremental benefit in the R-FC v. FC comparison is caused by the modelled advantage in OS between the arms. Given this, it is important to investigate to what extent the OS gain implicit in the model outputs can be supported by the evidence.

6.2.2. Investigation of structural uncertainty related to OS advantage between arms

Figure 7 shows the Kaplan Meier curves for the OS from the two arms of the CLL-8 trial data superimposed over the OS survival curves generated by the model in the R-FC v. FC comparison. It should be noted that the difference in the modelled OS survival curves is almost entirely due to the different rates of transfer from PFS to Progressed state between the arms of the model. Once patients have relapsed from the PFS to Progressed state, their probability of death in the model increases by over thirty times from 0.001196 and 0.001388 in the R-FC and FC arms respectively to 0.0405144 in both arms. It should be noted that although the base case model uses slightly different levels of mortality rate between arms for patients in the PFS state (based on the CLL-8 data) the impact of this difference on incremental OS is negligible.

Figure 7 : Kaplan Meier curves from CLL-8 trial and overall survival curves for R-FC v. FC generated by the economic model.



6.2.3. Additional Scenario Analysis – incremental OS removed between arms

Following a request from the ERG, Roche provided an analysis which removed any OS advantage between the arms of the model for the R-FC v. FC comparison. This was achieved by increasing the probability of death in the Progressed state in the R-FC arm of the model only by 315%. The results of this analysis are presented Table 31. The incremental utility shown is now derived solely from the utility advantage of being in the PFS state relative to the Progressed state.

Table 31 : Roche sensitivity analysis supplied to the ERG - 18 Dec 08 showing modelled effects of removing incremental OS in R-FC v. FC

Outcome measure	RFC	FC	Incremental
Mean Life Years (yrs)	4.66	4.65	0.00
Mean Life Years in PFS (yrs)	4.11	2.93	1.18
Mean life Years in Progression (yrs)	0.55	1.73	-1.18
Mean QALYs	3.62	3.38	0.24
Mean QALY in PFS	3.29	2.34	0.95
Mean QALY in Progression	0.33	1.04	-0.71

Cost-utility results	**RFC**	**FC**	**Incremental**
Mean Life Years (yrs)	4.66	4.65	0.00
Mean QALYs	3.62	3.38	0.24
Mean Total Cost	£21,204	£13,978	£7,226
Cost per Life Year Gained (£)			**£3,473,529**
Cost per QALY Gained (£)			**£30,336**

An alternative approach is to *decrease* the probability of death from the Progressed state for the *FC comparator arm only*. We have preferred this method in this scenario analysis since it results in an increased OS in the model, and the CLL-8 clinical trial data suggest a higher OS than that generated by the model base case (see Figure 7 above). This method, produces a very similar ICER output to the analysis provided by Roche shown in Table 31 above. We found a reduction of probability of death from the Progressed state *in the FC arm only* to 57.4% of the base case level has the effect of removing any OS gain between the two modelled arms. The outputs from this analysis are shown in Table 32.

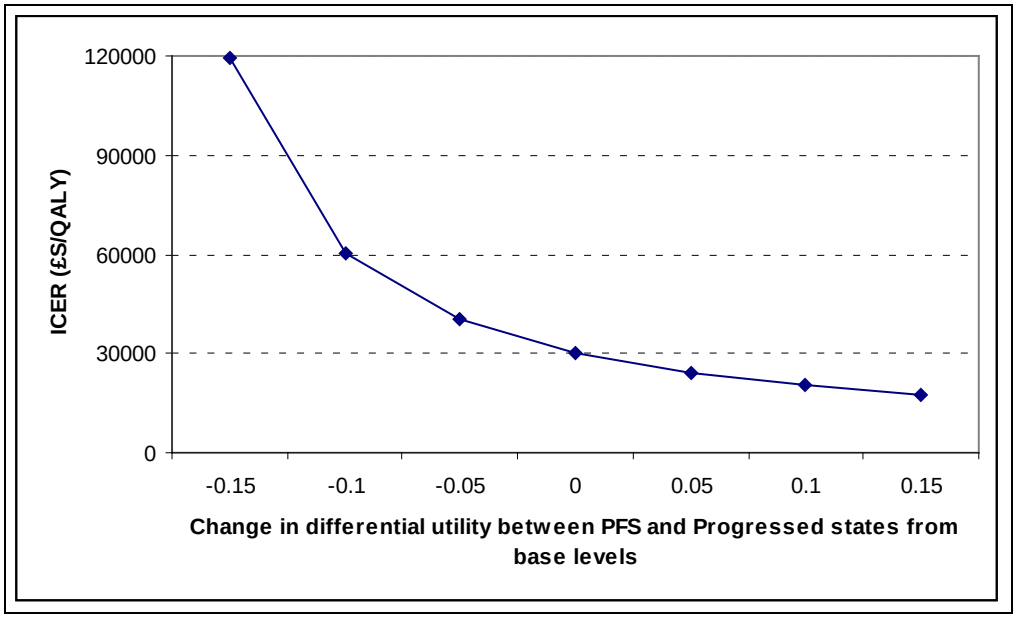
Table 32 : Main outputs from the Roche model for the R-FC v. FC comparison when the progressed to death probability is reduced to 57.4% of base case in FC arm only.

Outcome measure	**RFC**	**FC**	**Incremental**
Mean Life Years (yrs)	5.73	5.72	0.00
Mean Life Years in PFS (yrs)	4.11	2.93	1.18
Mean life Years in Progression (yrs)	1.62	2.80	-1.18
Mean QALYs	**4.26**	**4.02**	**0.24**
Mean QALY in PFS	3.29	2.34	0.95
Mean QALY in Progression	0.97	1.68	-0.71
Cost-utility results	**RFC**	**FC**	**Incremental**
Mean Life Years (yrs)	5.73	5.72	0.00
Mean QALYs	4.26	4.02	0.24
Mean Total Cost	£25,595	£18,367	£7,228
Cost per Life Year Gained (£)			**£2,756,887**
Cost per QALY Gained (£)			**£30,304**

A key observation in this analysis is that once any OS gain in the model is removed, the model becomes very sensitive to the utility difference between the PFS state and

the Progressed state. Figure 8 below shows the effect on the ICER value of varying the utility difference between the PFS and Progressed state once incremental OS advantage across the arms of the model is removed. These outputs were obtained by varying the base case utilities in both states by the same amount in steps of 0.025 to contract or expand the utility differential. From this it can be seen that a reduction of the differential by 0.1 between PFS and Progressed state (i.e. the differential is reduced from 0.2 to 0.1) causes the ICER to climb from a value of £30,336 per QALY to £60,302 per QALY in this scenario analysis (i.e. where the OS survival advantage is removed from the model).

Figure 8 : Threshold analysis showing relationship between the model ICER and utility differential between PFS and progressed state for model when OS advantage is removed.



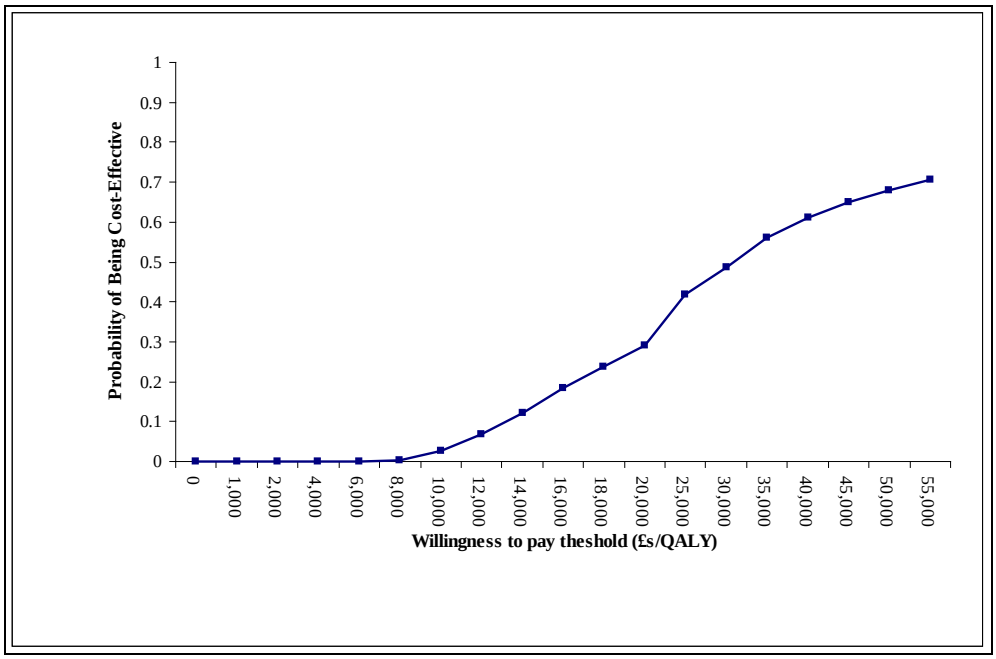
6.2.3.1. Probabilistic Analysis

In order to further explore the impact of the removal of incremental OS in this analysis we used the Roche model to carry out a series of probabilistic simulations to investigate uncertainty using PSA.

In the first instance, we ran the PSA simulation with the OS advantaged between arms removed from the model using the second method described above (i.e. where the probability of death in progressed state in the FC arm only is reduced to 57.4% of base case level). This generated the CEAC shown in Figure 9. This indicates a probability

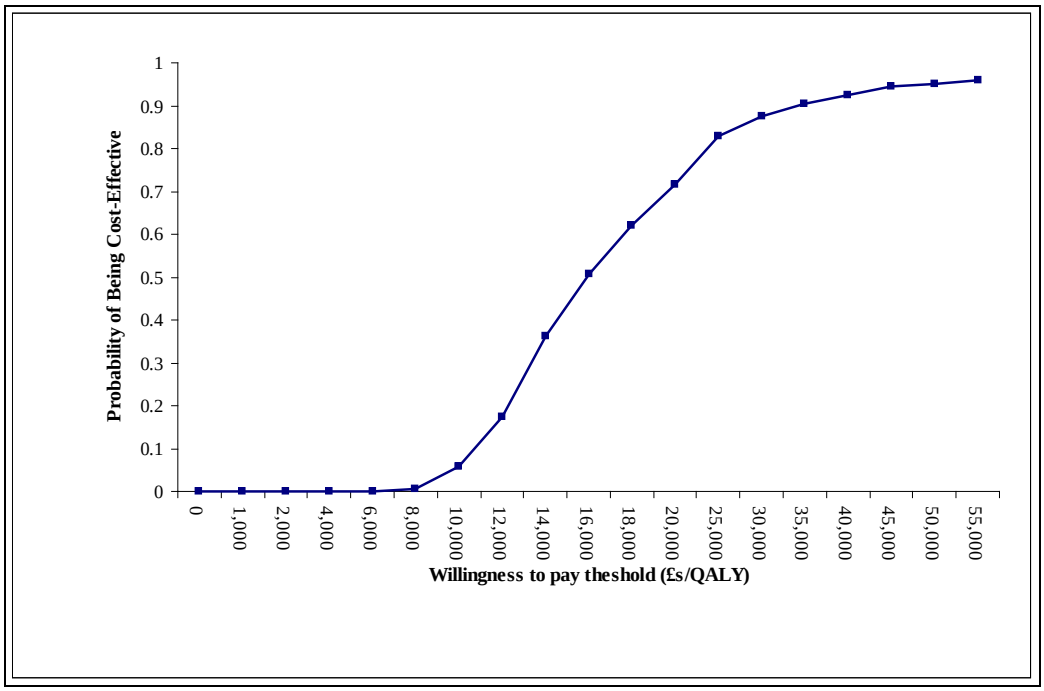
that the R-FC treatment is cost-effective relative to FC of 29% at the £20,000 willingness to pay threshold and a probability of 48.7% at the £30,000 threshold.

Figure 9 : Cost Effectiveness Acceptability Curve generated for the R-FC v. FC comparison for the scenario when OS advantage between arms is removed.



Secondly, we incorporated an additional sampled variable into the PSA which acted to scale the reduction in the transition probability from the Progressed to Death states in the FC arm only. For this we used a uniform distribution scaled between values of 1 and 0.574. The effect of this is to introduce into the model an underlying uncertainty about the size of the OS advantage across the R-FC and FC arms, but to assume that it lies somewhere between that represented by the Roche base case and the scenario shown above, where the OS advantage is removed. The results of this PSA are shown in Figure 10. This CEAC suggests a probability that R-FC is cost-effective at the £20,000 willingness to pay threshold of = 71.7% and a probability of = 87.6% at the £30,000 threshold.

Figure 10 : Cost Effectiveness Acceptability Curve generated for the R-FC v. FC comparison for the scenario when a probabilistic factor is used to scale the OS advantage between arms.



Whilst interpretation of the outputs presented in this scenario analysis is challenging, the results show clearly the importance of the structural assumptions inherent in the model relating to differential OS between arms in the model. This is against the background that, although the CLL-8 trial data show a significant advantage in PFS survival for R-FC treatment over FC alone, no advantage is statistically demonstrated for R-FC v FC for OS in this trial. This is consistent with other studies which have been unable to demonstrate a statistically significant OS advantage between treatments. Given the limited length of trial follow-up, there are obvious difficulties in demonstrating that an OS advantage exists empirically and a fuller examination of the issue would require an understanding of progression of the disease after first relapse. The fact that the model is for a first-line treatment with a relatively long period of survival post-progression makes this especially problematic.

The structural assumptions of the Roche base case analysis result in a significant modelled difference in OS between the R-FC and FC arms. A fuller understanding of the clinical background and natural progression of CLL, which go beyond the scope of this report, would be required to validate these assumptions.

7 Discussion

7.1. *Summary of clinical-effectiveness issues*

The literature search strategy was appropriate, replicable, and the hits appear correct in relation to the search date and databases/interfaces used. We are confident that there are no relevant and good quality studies which have not been presented in the submission.

The submission from Roche included one good quality RCT (CLL-8 trial) which compared to use of R-FC to FC alone. No direct head-to-head trials on the effects of R-FC versus chlorambucil were found. Therefore evidence for the effects of R-FC versus chlorambucil were presented via a mixed treatment comparison model (MTC). The CLL-8 trial included 5% of patients with Binet stage A at baseline, and an ECOG performance status of 0-2. Trial participants may therefore be slightly younger and healthier than those generally seen in practice. In the trial rituximab was taken in accordance with the dose that has been submitted for marketing authorisation and type II variation to EMEA and is presently anticipated will be the approved dose in the SmPC. However, it should be highlighted that the trial used a fully iv regimen, with both fludarabine (25 mg/ m2) and cyclophosphamide (250 mg/m2) being administered iv. Currently in the UK however, approximately 99% of all FC therapy is delivered orally, and whilst bio-equivalence between iv and oral regimens has been demonstrated, the addition of rituximab to this regimen would add an iv component to what is now a fully oral regimen.

Results from the CLL-8 trial highlighted a statistically significant treatment benefit for R-FC versus FC for PFS [median 39.8 months versus 32.2 months; HR 0.56 (95% CI: 0.43; 0.72)]. However, the initial treatment benefit for the R-FC regimen in terms of OS noted at the time of interim analysis (median follow-up time 20.7 months) was no longer maintained at slightly longer follow-up (median 25.4 months) [HR 0.72 (95% CI: 0.48; 1.09)]. Patients in the R-FC arm remained event free (disease progression, relapse, death or start of new CLL treatment) significantly longer than those in the FC-arm [39.8 months versus 31.1 months; HR 0.55 (95% CI: 0.43; 0.70)]. Response rates

also significantly favoured treatment with R-FC with 36.0% of patients in this arm achieving complete response, compared with 17.2% with FC. Partial response rates were not significantly different between trial arms at 50.1% and 55.5% respectively.

Grade 3 or 4 adverse events were higher in the R-FC arm (77%) compared to the FC arm (62%); mostly due to a higher incidence of blood and lymphatic system disorders (57% versus 41%). There were no differences in numbers of treatment discontinuations or the numbers of deaths considered related to therapy (2%).

Results from the mixed treatment comparison model, showed that in terms of both PFS and complete response rates treatment with R-FC was more favourable than with chlorambucil.

It is unclear whether the observed treatment benefit for use of rituximab combination therapy for PFS is associated with longer-term OS, and the magnitude of any such gain.

7.2. *Summary of cost-effectiveness issues*

The model presented by Roche uses a conventional Markov state-transition approach deploying the 3-states of PFS, Progressed and death. It is generally well constructed and presented.

Almost all data are based on the CLL-8 trial presented in the Roche submission. MTC analysis is used to parameterise effectiveness in the model for the R-FC v. Chlorambucil comparison.

An acknowledged limitation of the modelling analysis is the aggregation of all post relapse outcomes other than death into a single state, and the assumption that patients in this state can be treated as a single population with no distinction between model arms. The key impact of these limitations is the generation of incremental OS gain in the model although such OS advantage has not been demonstrated empirically.

Limited sensitivity analyses are used to explore the uncertainties generated by these assumptions and, specifically, no investigation is made of the uncertainty associated with differential levels of mortality from the progressed state between arms in this state. To address this limitation we used the Roche model to carry out a scenario

analysis which explores the outputs from the model when the OS advantage between arms is removed in the R-FC v. FC comparison. In this scenario it becomes clear that the assertion of an ICER below normally acceptable willingness-to-pay threshold becomes less robust.

The utility data used in the economic analysis are not drawn from an appropriate source. These values are explored in the sensitivity analysis and the model is shown to be insensitive to changes in the base case. However, this is because most of the incremental utility in the base case is generated by the difference in OS between arms, consequent on differential progression, rather than the difference in PFS per se. If this incremental OS is removed from the model then the analysis is highly sensitive to the differential utility between the PFS and progresses state (since this is the primary source of incremental utility).

References

(1) Catovsky D, Richards S, Matutes E, Oscier D, Dyer MJ, Bezares RF et al. Assessment of fludarabine plus cyclophosphamide for patients with chronic lymphocytic leukaemia (the LRF CLL4 Trial): a randomised controlled trial. Lancet 2007; 370(9583):230-239.

(2) Genactis CLL Monitor Q2 2008.

(3) Hallek M, Cheson BD, Catovsky D, Caligaris-Cappio F, Dighiero G, Dohner H et al. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. Blood 2008; 111(12):5446-5456.

(4) Rossi JF, van HA, de BK, Johnson SA, Bron D, Foussard C et al. Efficacy and safety of oral fludarabine phosphate in previously untreated patients with chronic lymphocytic leukemia. J Clin Oncol 2004; 22(7):1260-1267.

(5) Roche. CLL8 Final Clinical Study Report ML17102. 2008.

(6) Keating MJ, O'Brien S, Albitar M, Lerner S, Plunkett W, Giles F et al. Early results of a chemoimmunotherapy regimen of fludarabine, cyclophosphamide, and rituximab as initial therapy for chronic lymphocytic leukemia. J Clin Oncol 2005; 23(18):4079-4088.

(7) Tam CS, O'Brien S, Wierda W, Kantarjian H, Wen S, Do KA et al. Long-term results of the fludarabine, cyclophosphamide, and rituximab regimen as initial therapy of chronic lymphocytic leukemia. Blood 2008; 112(4):975-980.

(8) Byrd JC, Peterson BL, Morrison VA, Park K, Jacobson R, Hoke E et al. Randomized phase 2 study of fludarabine with concurrent versus sequential treatment with rituximab in symptomatic, untreated patients with B-cell chronic lymphocytic leukemia: results from Cancer and Leukemia Group B 9712 (CALGB 9712). Blood 2003; 101(1):6-14.

(9) Faderl S, et al. Update of experience with fludarabine, cyclophosphamide, mitoxantrone plus rituximab (FCM-R) in frontline therapy for chronic lymphocytic leukemia (CLL): a phase II study. Blood (ASH Annual Meeting Abstracts) Nov 2007[110], abstract 627. 2008.

(10) Kay NE, Geyer SM, Call TG, Shanafelt TD, Zent CS, Jelinek DF et al. Combination chemoimmunotherapy with pentostatin, cyclophosphamide, and rituximab shows significant clinical activity with low accompanying toxicity in previously untreated B chronic lymphocytic leukemia. Blood 2007; 109(2):405-411.

(11) Hillmen P, Skotnicki AB, Robak T, Jaksic B, Dmoszynska A, Wu J et al. Alemtuzumab compared with chlorambucil as first-line therapy for chronic lymphocytic leukemia. J Clin Oncol 2007; 25(35):5616-5623.

(12) Rai KR, Peterson BL, Appelbaum FR, Kolitz J, Elias L, Shepherd L et al. Fludarabine compared with chlorambucil as primary therapy for chronic lymphocytic leukemia. N Engl J Med 2000; 343(24):1750-1757.

(13) Knauf WU, Lissichkov T, Aldaoud A, et al. Bendamustine Versus Chlorambucil in Treatment-Naive Patients with B-Cell Chronic Lymphocytic Leukemia (B-CLL): Results of an International Phase III Study. Blood (ASH Annual Meeting Abstracts) Nov 2007[110], 2043. 2007.

(14) Eichhorst BF, Busch R, Stauch M, et al. No Significant Clinical Benefit of First Line Therapy with Fludarabine (F) in Comparison to Chlorambucil (Clb) in Elderly Patients (pts) with Advanced Chronic Lymphocytic Leukemia (CLL): Results of a Phase III Study of the German CLL Study Group (GCLLSG). Blood (ASH Annual Meeting Abstracts) Nov 2007[110], 629. 2007.

(15) O'Brien SM, Kantarjian H, Thomas DA, Giles FJ, Freireich EJ, Cortes J et al. Rituximab dose-escalation trial in chronic lymphocytic leukemia. J Clin Oncol 2001; 19(8):2165-2170.

(16) Chemoimmunotherapy with Fludarabine, Cycholphosphamide, and Rituximab versus Fludarabine and Cycholophamide improves response rates and progression-free survival in previously untreated patients with advanced chronic lymphocytic leukemia. Blood (ASH) Annual Meeting Abstracts; 2008.

(17) Flinn IW, Neuberg DS, Grever MR, Dewald GW, Bennett JM, Paietta EM et al. Phase III trial of fludarabine plus cyclophosphamide compared with fludarabine for patients with previously untreated chronic lymphocytic leukemia: US Intergroup Trial E2997. J Clin Oncol 2007; 25(7):793-798.

(18) Eichhorst BF, Busch R, Hopfinger G, Pasold R, Hensel M, Steinbrecher C et al. Fludarabine plus cyclophosphamide versus fludarabine alone in first-line therapy of younger patients with chronic lymphocytic leukemia. Blood 2006; 107(3):885-891.

(19) Jaksic B, Brugiatelli M, Krc I, Losonczi H, Holowiecki J, Planinc-Peraica A et al. High dose chlorambucil versus Binet's modified cyclophosphamide, doxorubicin, vincristine, and prednisone regimen in the treatment of patients with advanced B-cell chronic lymphocytic leukemia. Results of an international multicenter randomized trial. International Society for Chemo-Immunotherapy, Vienna. Cancer 1997; 79(11):2107-2114.

(20) Jaksic B, Brugiatelli M, Suciu S, et al. Fludarabine vs. high dose chlorambucil in advanced B-cell chronic lymphocytic leukemia (B-CLL): Results of EORTC phase-II randomized trial after 4 years of median follow-up. Blood 2000; November 16, 96(11 PT. 1):758a.

(21) Cladribine (CdA) or Fludarabine (F) or High-Dose Intermittent Chlorambucil (Chl) as First-Line Treatment of Symptomatic Chronic Lymphocytic Leukemia? First Interim Analysis of Data from the International Randomized Phase III Trial. American Society of Hematology 46th Conference; 2004.

(22) Leporrier M, Chevret S, Cazin B, et al. Randomized comparison of fludarabine, CAP, and ChOP in 938 previously untreated stage B and C chronic lymphocytic leukemia patients. Blood 2001; 15 October 98(8):2319-2325.

(23) CLL Trialists' Collaborative Group. Chemotherapeutic options in chronic lymphocytic leukaemia. J Natl Cancer Inst 1999; 91:861-868.

(24) O'Brien S, Kantarjian H, Thomas D, Giles FJ, Freireich EJ, Cortes J. Rituximab dose-escalation trial in chronic lymphocytic leukemia. J Clin Oncol 2001; 19(8):2165-2170.

(25) Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials 1996; 17(1):1-12.

(26) Drummond M, O'Brien B, Stoddart G, Torrance G. Methods for the economic evaluation of health care programmes
47. Oxford, UK: Oxford University Press; 1997.

(27) Schulz KF, Chalmers I, Hayes RJ, Altman D. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA 1995; 273:408-12.

(28) Moher D, Pham B, Jones A. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analysis? Lancet 1998; 352:609-13.

(29) Spitzer WDAHJ. Measuring the quality of life of cancer patients: a concise QL-Index for use by physicians. Journal of chronic diseases 1981;585-597.

(30) Aaronson NK, Ahmedzai S, Bergan B. The European Organization for Research and Treatment of Cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993; 85:365-76.

(31) 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):683-691.

(32) Hancock S, Wake B, Hyde C. Fludarabine as first line therapy for Chronic Lymphocytic Leukaemia. 2002. A West Midlands Health Technology Assessment Collaboration.

(33) Holzner B, Kemmler G, Sperner-Unterweger B, Kopp M, Dunder M, Mergreiter R. Quality of life measurements in oncology - a matter of the measurement instrument? European Journal of Cancer 2001; 37:2349-2356.

(34) National Institute for Health and Clinical Excellence. Guide to the Methods of Technology Appraisal. National Institute for Health and Clinical Excellence website [2008 Available from:
URL:http://www.nice.org.uk/aboutnice/howwework/devnicetech/technologyappraisalprocessguides/guidetothemethodsoftechnologyappraisal.jsp?domedia=1&mid=B52851A3-19B9-E0B5-D48284D172BD8459

(35) Philips Z, Bojke L, Sculpher M, Claxton K, Golder S. Good Practice Guidelines for Decision-Analytic Modelling in Health Technology Assessment: A Review and Consolidation of Quality Assessment. Pharmacoeconomics 2006; 24(4):355-371.

(36) Watson L, Wyld P, Catovsky D. Disease burden of chronic lymphocytic leukaemia within the European Union. Eur J Haematol 2008; 81(4):253-258.

(37) Hancock S, Wake B, Hyde C. Fludarabine as first line therapy for Chronic Lymphocytic Leukaemia. A West Midland Health Technology Assessment Collaboration. 2002.

(38) Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. American Journal of Clinical Oncology 5(6):649-55, 1982.

Appendix 1: Binet staging system in Chronic Lymphocytic Leukaemia

	Features	% of patients
Binet Stage		
A	**<3 lymphoid areas involved**	**60**
B	**>3 lymphoid areas involved**	**30**
C	**Haemoglobin <10g/dl or platelets , 100 X 10^9/L**	**10**

Appendix 2: ECOG performance status criteria

Grade	ECOG
0	Fully active, able to carry on all pre-disease performance without restriction
1	Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g., light house work, office work
2	Ambulatory and capable of all selfcare but unable to carry out any work activities. Up and about more than 50% of waking hours
3	Capable of only limited selfcare, confined to bed or chair more than 50% of waking hours
4	Completely disabled. Cannot carry on any selfcare. Totally confined to bed or chair
5	Dead

Source: Eastern Cooperative Oncology Group, Robert Comis M.D., Group Chair. [38]

Appendix 3: Supporting Phase II Studies

Study: Keating et al (2005)19, Tam et al (2008)20 **R-FC as initial treatment for CLL**	
Rationale and Purpose	To test the efficacy and tolerability of adding rituximab to the combination of fludarabine and cyclophosphamide for the initial treatment of CLL, with the hope of increasing CR rates to greater than 50%.
Design	Single-arm, open label Phase II study of 300 patients.
Participants	300 patients aged 18 years or older with previously untreated CLL requiring therapy as indicated by NCI 1996 guidelines. Median age was 57, with 14% being 70 or older. 61% of patients had Rai stage I-II disease, and 36% had Rai stage III-IV disease. Cytogenetics via conventional karyotyping was available for 222 patients (FISH was not available); of these 30% had clonal abnormalities and 4% had abnormalities involving chromosome 17. FISH, IgvH and ZAP-70 were not clinically available at the time of study recruitment.
Interventions	6 cycles of R-FC given every 28 days. All medication given i.v. Rituximab : 375mg/m2 cycle 1, 500mg/m2 cycles 2-6; Fludarabine 25-30mg/m2 for 3 days each course and cyclophosphamide 250-300mg/m2 for 3 days each course.
Outcomes	Responses as according to NCI criteria, overall survival, failure-free survival, time to progression.
Added comments	A historical comparison of R-FC against previous frontline fludarabine-based regimes at the MD Anderson Cancer Center is also presented.
Study: Kay et al, 2007i **R-PC as initial treatment for CLL**	
Rationale and Purpose	To evaluate the efficacy and tolerability of pentostatin in combination with cyclophosphamide and rituximab for the initial treatment of CLL.
Design	A Phase II, single arm non-randomised study.
Participants	65 patients with progressive CLL requiring treatment for the first time, as defined by NCI criteria. Median age was 63 years, 34/64 patients (53%) had high risk disease according to Rai stage (i.e. stage 3 or 4 disease), 71% had unmutated IgVH, 34% were CD38-positive, and 28% were ZAP-70-positive. Thirty patients (47%) had one anomaly detected by fluorescence in situ (FISH) hybridisation, and 21

	(33%) had complex FISH defects.
Interventions	6 cycles of treatment, given every 21 days consisting of Pentostatin 2mg/m2, cyclophosphamide 600mg/m2, and rituximab 375mg/m2 all given on day 1 of cycles 2-6. In cycle 1 , more doses of rituximab were given, with 100mg/m2 on day 1 and 375mg/m2 on days 3 and 5.
Outcomes	Responses graded according to NCI criteria, and all patients followed up for 5 years or until disease progression. Minimum residual disease was also tested for in responding patients.
Added comments	Pentostatin is a purine analogue similar to fludarabine but is also a potent transition state inhibitor of the enzyme adenosine deaminase (ADA). This inhibition, as well as direct inhibition of RNA synthesis and increased DNA damage contributes to the overall cytotoxic effect of pentostatin.
Study: Faderl et al, 2007ii **R-FCM as initial treatment for CLL**	
Rationale and Purpose	To evaluate the efficacy and tolerability of fludarabine, cyclophosphamide, mitoxantrone and rituximab for the initial treatment of CLL
Design	A Phase II, single arm non-randomised study.
Participants	30 patients with progressive CLL requiring treatment, as defined by NCI criteria. The median age was 57 years (range 38-69). Fourteen patients (48%) were male. Four patients (14%) had Rai stage ≥ 3 disease. Median β2-microglobulin was 2.6 mg/L (range 1.4-4) and the median WBC was 59.9 x 109/L (range 5.6-355). Two patients had 11q23 and 17p- abnormalities by cytogenetics/fluorescent in situ hybridisation (FISH). Unmutated IgVH was reported in 12/17 patients (71%) and ZAP-70 was positive in 11/19 patients (58%).
Interventions	Fludarabine (25 mg/m2 i.v. d 2-4), cyclophosphamide (250 mg/m2 i.v. d 2-4), mitoxantrone (6 mg/m2 i.v. d 2), rituximab (375 mg/m2 i.v. d 1) and pegfilgrastim (6 mg s.c. d 4) in cycle 1. For cycles 2 to 6, FCM started on day 1 together with 500 mg/m2 of rituximab followed by pegfilgrastim on day 3. Cycles were repeated every 4 to 6 weeks.
Outcomes	Responses graded according to NCI criteria, at 3 and 6 months post starting therapy.
Added comments	Further publication on this study expected next year.
Study: Byrd et al, 200329 and 2005iii Combining Rituximab and Fludarabine for the initial treatment for CLL	
Rationale and	To investigate the efficacy, tolerability and optimal schedule of

Purpose	administration of rituximab with fludarabine in previously untreated patients. Two schedules of rituximab were presented: concurrent, or sequential.
Design	Non-comparative, randomised Phase II study.
Participants	104 patients with 51 in Arm A (concurrent), and 53 in Arm B (sequential). All required treatment for CLL as defined by NCI criteria. Median age was 64. 59% of patients had Rai Stage I-II disease, the rest had Stage III-IV.
Interventions	Concurrent Regime: 6 monthly courses of fludarabine (25mg/m2 iv for 5 days) concurrently with rituximab (375mg/m2 – 2 doses in cycle one and one in each of the subsequent 5), followed 2 months later by 4 weekly doses of rituximab (375mg/m2) for consolidation therapy. Sequential Regime: Treatment with fludarabine alone followed by rituximab consolidation therapy 2 months later, doses as described above. Consolidation therapy was only given to those who responded to initial treatment.
Outcomes	Response rates, progression-free and overall survival were all measured for both concurrent and sequential groups. Please note that this was study was not designed to compare both arms and contained rituximab in both arms, and is therefore not appropriate for discussion in section in 6.3.
Added comments	The 2004 publication compared all the pooled results of CALGB 9712 versus an older CALGB study (9011) of fludarabine monotherapy in a similar population.

[i] Kay NE et al. Combination chemoimmunotherapy with pentostatin, cyclophosphamide, and rituximab shows significant clinical activity with low accompanying toxicity in previously untreated B chronic lymphocytic leukemia. Blood.2007:109(2);405–411

[ii] Faderl S et al. Update of experience with fludarabine, cyclophosphamide, mitoxantrone plus rituximab (FCM-R) in frontline therapy for chronic lymphocytic leukemia (CLL): a phase II study. Blood (ASH annual meeting abstracts). 2007;110:Abstract 627

[iii] Byrd JC et al. Addition of rituximab to fludarabine may prolong progression-free survival and overall survival in patients with previously untreated chronic lymphocytic leukemia: an updated retrospective comparative analysis of CALGB 9712 and CALGB 9011. Blood 2005;105(1):49-53