

Rituximab for the First-Line Maintenance Treatment of Follicular Non-Hodgkin's Lymphoma

A NICE Single Technology Appraisal

Janette Greenhalgh · Adrian Bagust · Angela Boland ·
Michaela Blundell · James Oyee · Sophie Beale · Yenal Dundar ·
Juliet Hockenhull · Chris Proudlove · Patrick Chu

Published online: 11 April 2013

© The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract The National Institute for Health and Clinical Excellence (NICE) invited the manufacturer of rituximab (RTX) [Roche] to submit evidence for the clinical and cost effectiveness of RTX as first-line maintenance treatment for patients with follicular non-Hodgkin's lymphoma (fNHL) whose disease has responded to induction therapy with RTX plus cytotoxic chemotherapy (R-CTX) in accordance with the Institute's Single Technology Appraisal (STA) process. The Liverpool Reviews and Implementation Group (LRiG) at the University of Liverpool was commissioned to act as the Evidence Review Group (ERG). This article summarizes the ERG's review of the evidence submitted by the manufacturer and provides a summary of the Appraisal Committee's (AC) decision. The clinical evidence was derived from a multi-centred, open-label, randomized phase III study (PRIMA) comparing first-line maintenance treatment with RTX with observation only in 1,018 patients with previously untreated advanced fNHL. Median time to event (MTE) for the primary endpoint of progression-free survival (PFS) in the RTX arm was not estimable due to data immaturity; median PFS in the observation arm was 48.36 months. A statistically significant benefit of RTX maintenance therapy for PFS was reported (hazard ratio [HR] 0.55, 95 % CI 0.44–0.68; $p < 0.0001$). Statistically significant differences in favour of RTX were also

reported for a range of secondary endpoints. Assessment of overall survival benefit could be not made due to insufficient events. The ERG's main concern with the clinical-effectiveness data presented was their lack of maturity. The submitted incremental cost-effectiveness ratio was within the NICE threshold. The ERG questioned the model on a number of grounds, particularly the use of Markov methodology rather than patient simulations, the impact of patient age on the outcome and the projective PFS modelling. The ERG considered it impossible to draw firm conclusions regarding the clinical or cost effectiveness of the intervention as the dataset was as yet too immature. At a third meeting, the AC concluded that RTX could be recommended as first-line maintenance treatment for patients with fNHL whose disease has responded to induction R-CTX.

Key Points for Decision Makers

- The clinical and cost effectiveness evidence presented was based on a well-designed randomized controlled trial (RCT). The clinical data for the effectiveness of rituximab as a maintenance treatment appeared to be promising. However, trials with immature outcome data pose problems in the assessment of longer-term clinical and cost effectiveness; this is particularly the case for diseases that are of prolonged duration. These issues were clearly apparent during this appraisal
- Data from trials that are stopped early based on a beneficial treatment effect (typically indicated by a statistically significant hazard ratio) should be viewed with caution as the efficacy may be overestimated
- In diseases that are of prolonged duration, it may be more useful to model outcomes using simulation based on patient-level data rather than employing a Markov approach

J. Greenhalgh (✉) · A. Bagust · A. Boland · M. Blundell ·
J. Oyee · S. Beale · Y. Dundar · J. Hockenhull
Liverpool Reviews and Implementation Group,
University of Liverpool, 2nd Floor, Whelan Building,
The Quadrangle, Brownlow Hill, Liverpool L69 3 GB, UK
e-mail: janette.greenhalgh@liverpool.ac.uk

C. Proudlove
North West Medicines Information Centre, Liverpool, UK

P. Chu
Royal Liverpool Hospital, Liverpool, UK

1 Introduction

The National Institute for Health and Clinical Excellence (NICE) is an independent organization responsible for providing national guidance to the NHS in England and Wales on a range of clinical and public health issues, as well as appraisal of new health technologies. The NICE Single Technology Appraisal (STA) process [1] is specifically designed for the appraisal of a single health technology for a single indication, where most of the relevant evidence lies with one manufacturer or sponsor. Typically, the process is used for new pharmaceutical products close to launch. The evidence for an STA is principally derived from a submission by the manufacturer/sponsor of the technology, which should be based on a specification developed by NICE. The manufacturer's submission (MS) is critiqued by members of the independent Evidence Review Group (ERG) who produce a report to be considered by the NICE Appraisal Committee (AC).

The NICE AC considers the submissions from the manufacturer and the ERG alongside testimony from experts, patients and other stakeholders to formulate preliminary guidance. All stakeholders have an opportunity to comment on this preliminary guidance and the manufacturer has the opportunity to provide a supplementary evidence submission. The AC meets again to produce the final guidance (final appraisal determination [FAD]). This article presents a summary of the ERG report for the STA of rituximab (RTX) as first-line maintenance treatment for patients with follicular non-Hodgkin's lymphoma (fNHL) whose disease has responded to induction therapy with RTX plus cytotoxic chemotherapy (R-CTX). This is one in a series of STA summaries being published in *Pharmacoeconomics* [2–22].

Full details of all the relevant appraisal documents (including the appraisal scope, ERG report, manufacturer and consultee submissions, Appraisal Consultation Document [ACD], FAD and comments on each of these) can be found on the NICE website [23].

2 The Decision Problem

fNHL is one of a group of diseases known collectively as non-Hodgkin's lymphomas (NHLs), cancers arising from the lymphoid cells of the immune system. These cells normally have a key role in protecting the body from pathogenic microorganisms. Malignant transformation of lymphocytes results in their uncontrolled replication usually starting within the lymph nodes, mainly those of the neck, armpits and groin. Swelling of these structures often provides the first clinical manifestation of illness, although other symptoms including fever, drenching night sweats,

weight loss ('B-symptoms') and tiredness may also be present at diagnosis or develop later. The incidence of NHL is similar in men and women and increases with age: rates increase sharply in people over 50 and around two-thirds of all cases are diagnosed in people over 60 years of age [24].

The estimated number of new cases of NHL in the USA for 2011 is 66,000 [25]. Australian projections for new cases of NHL for 2010 are 4,000 [26]. For Europe as a whole, the estimated number of new cases for 2008 was 52,000 [27]. In England and Wales in 2008, there were almost 12,000 new cases of NHL recorded [24]. Between 22 % and 40 % of NHLs are fNHL [28, 29]. The majority of cases of fNHL are diagnosed at an advanced stage (III/IV) [30].

Typical estimates of median survival for patients with advanced fNHL are 8–10 years from diagnosis [31, 32]. However, recent evidence [33–36] suggests that survival may be longer, probably as a consequence of improved treatment. The natural course of advanced disease is of a series of treatment-induced remission followed by relapse, with each remission duration becoming shorter and fewer patients responding to each cycle of therapy [37, 38].

Patients in the UK diagnosed with advanced fNHL and who are symptomatic, have a high tumour burden, rapid disease progression or other key features are normally offered induction treatment with R-CTX [29].

The aim of using RTX maintenance therapy (375 mg/m² body surface area, by intravenous infusion once every 2 months until disease progression, or for a maximum of 2 years) following response to induction therapy is to extend and deepen the first remission [37, 38]. It would be expected, as in the relapsed/refractory setting, that extended progression-free survival (PFS) as a consequence of RTX maintenance therapy will delay the time to first relapse and, therefore, the burden of further CTX treatment [39, 40].

NICE developed a scope for the assessment of RTX, which specified that the clinical and cost effectiveness of this drug should be established within its licensed indication relative to standard management without RTX maintenance therapy (observation) and ibritumomab tiuxetan (Zevalin). Five outcome measures were considered relevant to the appraisal: PFS, overall survival (OS), response rates, adverse effects (AEs) of treatment and health-related quality of life (HR-QOL). The time horizon of analysis was stipulated as sufficiently long to reflect any differences in costs or outcomes between the technologies being compared.

At the time of this appraisal, the licence for RTX as a maintenance therapy was limited to use in patients with relapsed/refractory lymphoma. However, the European Medicines Agency's (EMA) Committee on Medicinal

Products for Human Use had issued a positive opinion to extend the use of first-line RTX maintenance therapy to include the treatment of fNHL patients responding to induction therapy. The licence extension was granted by the EMA in December 2010 [41].

3 Independent ERG Report

The ERG examined and critiqued the initial MS and two subsequent supplementary evidence submissions from the manufacturer as well as taking into consideration the manufacturer's response to the ERG's request for clarification on a number of issues. The ERG report comprised a critical review of the evidence for the clinical and cost effectiveness of the technology and embodied three aims:

- To assess whether the MS conformed to the methodological guidelines issued by NICE
- To assess whether the manufacturer's interpretation and analysis of the evidence were appropriate
- To indicate the presence of other sources of evidence or alternative interpretations of the evidence that could help inform NICE guidance

In addition to providing this detailed critique, the ERG modified a number of key assumptions and parameters within the manufacturer's economic model to examine the impact of such changes. The next section summarizes the evidence submitted by the manufacturer and the ERG's review of that evidence.

3.1 Clinical Evidence

The clinical-effectiveness evidence was derived from a single trial, PRIMA (Primary Rituximab and Maintenance trial) [42], which was unpublished at the time the MS was submitted to NICE. The PRIMA trial results are now published [43]. The MS was based on data from the clinical study report. The PRIMA trial [42] was an international, multi-centred, open-label, randomized phase III study ($n = 1018$) designed to evaluate the efficacy of RTX maintenance treatment compared with observation for patients with advanced fNHL whose disease had responded to induction treatment with R-CTX.

The outcomes of the trial are described in Table 1. For the primary endpoint of PFS (defined as first documented disease progression, relapse or death from any cause), at median follow-up of 38 months, the median time to event (MTE) for patients in the RTX arm was not estimable due to insufficient PFS events; the median PFS in the observation arm was 48.36 months. A statistically significant benefit of RTX maintenance therapy for PFS was reported (hazard ratio [HR] 0.55, 95 % CI 0.44–0.68; $p < 0.0001$).

For secondary endpoints at a median follow-up of 36 months, statistically significant differences were reported for event-free survival (HR 0.59, 95 % CI 0.48–0.72; $p < 0.0001$), time to next anti-lymphoma treatment (HR 0.60, 95 % CI 0.47–0.76; $p < 0.0001$), time to next CTX treatment (HR 0.62, 95 % CI 0.47–0.81; $p = 0.0005$) and overall response rate (difference in proportions 17.98, 95 % CI 12.3–23.6; $p < 0.0001$), although the only calculable MTE was for event-free survival (first documented progression, relapse, initiation of new anti-lymphoma treatment or death from any cause) in the observation arm. No meaningful conclusions regarding OS could be drawn due to the lack of deaths. The incidence of AEs was higher in the RTX arm than the observation arm; however, no unexpected safety findings were reported. No statistically significant differences were reported on the outcome of QOL as measured by the Functional Assessment of Cancer Therapy—General (FACT-G) questionnaire [44] and the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC-QLQ C30) [45].

The manufacturer dismissed Zevalin as a comparator on the grounds that (1) there is no clinical evidence to support the clinical benefit of Zevalin in patients with untreated advanced fNHL, and (2) there is minimal Zevalin use in UK clinical practice.

3.1.1 Critique of the Clinical Evidence and Interpretation

The ERG considered the PRIMA trial [42] to be large and well designed, with safeguards to mitigate against possible bias in monitoring and assessment (particularly important since the trial was open label). The ERG's opinion was that the results of the trial were generalizable to UK clinical practice as the participants in the trial were comparable with patients seen in UK clinical practice and the majority were recruited from countries with similar care pathways to the UK. The ERG considered that the manufacturer had made a convincing case for not including Zevalin as a comparator.

The ERG's chief concern was the immaturity of the outcome data. This immaturity stemmed from two factors: the early closure of the trial (recommended by the study Data and Safety Monitoring Committee on the grounds of benefit [i.e. that the primary endpoint of PFS was crossed in favour of the maintenance arm]) and the limited follow-up data available for report in the MS. At the time of study closure, 1,018 patients were enrolled, 505 and 513 in the RTX and observation arms, respectively. Of these patients, only 93 (18.4 %) in the RTX arm and 174 (33.9 %) in the observation arm had progressed at the time of analysis. None of the patients had been followed up for more than 4 years.

Table 1 Key outcomes of the PRIMA trial

Endpoint	RTX (<i>N</i> = 505)	Observation (<i>N</i> = 513)	HR (95 % CI)	<i>P</i> value
<i>Primary: investigator-assessed PFS (38 months)</i>				
Median time to event (95 % CI)	NE	48.36 (42.09 to NS)	0.55 (0.44–0.68)	<0.0001
<i>Secondary endpoints (36 months)</i>				
Event-free survival				
Median time to event	NE	1381 days [45.4 months]		
25th percentile	1000 days [32.8 months]	497 days [16.3 months]	0.59 (0.48–0.72)	<0.0001
1-year event-free rate (95 % CI)	0.89 (0.86–0.92)	0.81 (0.78–0.84)		
Overall survival				
Median time to event	NE	NE		
25th percentile	NE	NE	0.87 (0.45–1.47)	0.6010
1-year event-free rate (95 % CI)	0.99 (0.98–1.00)	0.99 (0.98–1.00)		
Time to next anti-lymphoma treatment				
Median time to event	NE	NE		
25th percentile	1328 days [43.6 months]	775 days [25.5 months]	0.60 (0.47–0.76)	<0.0001
1-year event-free rate (95 % CI)	0.92 (0.89–0.94)	0.89 (0.87–0.92)		
Time to next CTX treatment				
Median time to event	NE	NE		
25th percentile	NE ^a	975 days [32.0 months]	0.62 (0.47–0.81)	0.0005
1-year event-free rate (95 % CI)	0.92 (0.90–0.95)	0.91 (0.89–0.94)		

The table is adapted from the manufacturer's submission and is reproduced from Greenhalgh et al. [54] (Crown copyright)

P values and HRs were calculated using the stratified log-rank test and stratified Cox regression for time-to-event endpoints, respectively. Stratification factors were induction treatment received and response to induction treatment. *P* values for response rate were calculated using the χ^2 test

CI confidence interval, CTX cytotoxic chemotherapy, ERG Evidence Review Group, HR hazard ratio, NE not estimable, NS not stated, PFS progression-free survival, RTX rituximab

^a The ERG notes that this NE is likely to be a typing error as a figure was reported at an earlier time point

The ERG's concerns regarding the immaturity of the data were supported by the findings of a meta-analysis [46], which reported large differences in treatment effect sizes between trials that were stopped early and similar trials that ran their full course. This finding, based on a comparison of 91 trials (cancer and non-cancer trials) that were terminated early based on a beneficial treatment effect (typically indicated by a statistically significant hazard ratio) with 424 similar trials that ran to full term, was robust regardless of the methodological quality of the trials or the presence of statistical stopping rules [46].

3.1.2 ERG Conclusions on the Submitted Clinical Evidence

Whilst the clinical trial data appeared to show a benefit for patients in the RTX arm, the ERG considered that the data submitted were too immature to draw definite conclusions. Longer follow-up with more events would be required before informed decisions could be made regarding the clinical effectiveness of RTX as a first-line maintenance therapy compared with observation.

3.2 Cost-Effectiveness Evidence

The MS included a review of the literature of the cost effectiveness of RTX as a first-line maintenance treatment in fNHL; no relevant published economic evaluations were identified.

The manufacturer undertook a de novo economic evaluation of RTX first-line maintenance therapy (intervention) compared with observation (comparator) for the treatment of patients with fNHL responding to first-line induction. The economic model was developed over a 25-year time horizon to capture the lifetime costs and QALYs of an average patient with fNHL and took an NHS and Personal Social Services perspective.

The economic model developed by the manufacturer was a four-state Markov model with a cycle length of 1 month. Patients enter the model in the PFS1 state (disease-free) having successfully completed induction therapy with R-CTX. Patients who responded to R-CTX either receive RTX monotherapy maintenance therapy (for 2 years) in the intervention arm or do not receive any treatment in the comparator arm. After initial treatment, at

the end of each cycle patients remain in PFS1 or progress to PFS2 (remission/full response to second-line RTX monotherapy in accordance with NICE's technology appraisal [TA] 137 [47]) or die. Once in the PFS2 health state, a patient may remain in this health state, die at the end of each cycle or move to progressive disease (PD). Patients in the PD state cannot move back to PFS2; they can either remain in PD or die at the end of each cycle. Death is an absorbing health state.

The model was populated with clinical-effectiveness results from the PRIMA trial [42] at median follow-up of 38 months and from an earlier trial of maintenance treatment with RTX following second-line CTX treatment, known as EORTC 20981 [39, 40]. HR-QOL utility values were taken from a paper [48] reporting a study that had originally been commissioned by Roche, and the main sources of cost data were the *NHS Reference Cost Schedule 2008–2009* [49] and the *British National Formulary* (56–59) [50].

The manufacturer's base-case incremental cost-effectiveness ratio (ICER) for RTX for the first-line maintenance treatment of patients with fNHL compared with observation was £15,978 per QALY gained and £14,697 per life-year gained. The manufacturer was confident that probabilistic sensitivity analysis results supported the robustness of the cost effectiveness of RTX maintenance therapy compared with observation in patients with fNHL responding to first-line induction therapy.

3.2.1 Critique of the Cost-Effectiveness Evidence and Interpretation

The ERG identified that the main source of uncertainty in the economic model was the immature nature of the primary data source, i.e. the PRIMA trial [42]. The submitted model projected future benefits in terms of increased patient time in PFS, and this was the dominant driver of cost effectiveness. In the manufacturer's base-case results, the model estimated the mean (undiscounted) survival for patients without maintenance therapy as 11.44 years, and 13.38 years for those with RTX maintenance therapy: a gain of 1.94 years. The model also estimated the mean time spent in PFS as 8.64 years (observation only) and 10.65 years (RTX maintenance): a gain of 2.01 years. Virtually all this benefit was generated in the first PFS phase of the model, with the implication that PFS gains are translated almost entirely (96.6 %) into OS gains. The ERG considered this a 'best possible' scenario that required robust supportive evidence from clinical trials before it could be accepted. At present [51], there is no unequivocal evidence from any clinical trial or meta-analysis that RTX maintenance treatment of patients with fNHL offers any significant OS gains, despite good

evidence of PFS gains. The PFS advantage from first-line RTX maintenance therapy directly measurable from the mature PRIMA trial data [42] (up to 800 days from randomization using the latest clinical data available) amounted to no more than 60 days.

A number of issues relating to the model design and structure, and to the implementation of the model were also raised by the ERG. These included:

- *Health states and pathways:* Due to the long/lifetime duration and complex pathways associated with fNHL, it is very difficult to represent real-life patterns of care adequately within a Markov framework as Markov models require the definition of essentially homogeneous health states in which patients share common risks, utility and treatment costs regardless of their prior history. There is a strong rationale for employing patient-level simulation methods in preference to a Markov model in such situations.
- *Principal model driver:* PFS is the principal driver of health gain in the model. However, the ERG considered 'time to next anti-lymphoma treatment' to be a more appropriate outcome variable as this provides a better reflection of clinical reality.
- *Use of both PRIMA [42] and EORTC 20981 [39, 40] trial data to populate the model:* The populations in these two trials are at different stages of their disease career and therefore the appropriateness of using data from the two trials is questionable.

The main implementation issue identified by the ERG was the manufacturer's use of different utility values for PFS1 (disease-free) and PFS2 (remission/full response to therapy); there should be no difference in utility as both groups of patients are in 'remission/full response'. Other identified issues included inconsistent logic in estimating deaths, errors in the calculations of patients eligible for second-line treatment and inaccurate and inappropriate estimation of event rates from EORTC 20981 [39, 40]. Errors were also noted regarding discounting, mid-cycle correction, AE costs, costs of RTX treatment, timing of RTX doses and health state costs.

3.2.1.1 ERG Sensitivity Analyses The ERG carried out sensitivity analyses (SAs) around two aspects of the manufacturer's model that appeared to be influential in governing the cost-effectiveness results; namely, patient age and the assumed duration of clinical benefit (Table 2).

The ERG's clinical advisor confirmed that the age at which a patient is diagnosed with fNHL and begins their first course of CTX is of great importance in determining the potential benefit that may accrue from the use of a novel regimen. The ERG requested additional results via the clarification process to allow a comparison of clinical

Table 2 ERG exemplification of results with mean patient age adjusted to 62.5 years and using the reported PFS hazard ratio of 0.55 (PRIMA trial original)

Model scenario/alteration	Cost per patient		QALYs per patient		Increment		ICER
	RTX maintenance	Observation	RTX maintenance	Observation	Cost	QALYs	
ERG revised base case	£70,666	£52,823	7.871	6.830	£17,843	1.041	£17,136
28-months' effect							
70 % PFS converts to OS	£66,227	£48,983	6.829	6.437	£17,244	0.392	£43,934
80 % PFS converts to OS	£66,607	£48,983	6.869	6.437	£17,624	0.432	£40,822
90 % PFS converts to OS	£66,986	£48,983	6.908	6.437	£18,003	0.471	£38,234
36-months' effect							
70 % PFS converts to OS	£65,740	£48,983	6.911	6.437	£16,756	0.474	£35,327
80 % PFS converts to OS	£66,198	£48,983	6.959	6.437	£17,215	0.522	£33,000
90 % PFS converts to OS	£66,655	£48,983	7.006	6.437	£17,672	0.569	£31,067
48-months' effect							
70 % PFS converts to OS	£65,098	£48,983	7.022	6.437	£16,114	0.585	£27,558
80 % PFS converts to OS	£65,662	£48,983	7.080	6.437	£16,679	0.643	£25,939
90 % PFS converts to OS	£66,224	£48,983	7.138	6.437	£17,291	0.701	£24,595

Reproduced from Liverpool Reviews and Implementation Group [55]

ERG Evidence Review Group, ICER incremental cost-effectiveness ratio, OS overall survival, PFS progression-free survival, RTX rituximab

effectiveness between three age-based subgroups. Despite the immaturity of the PRIMA trial data [42] and the unsophisticated nature of the analysis, there appeared to be an emerging trend indicating a reduction in clinical effect as patient age increased—a curvilinear trend in odds ratio (OR), equivalent to a linear trend in relative risk. The HR of PFS in the base-case model was adjusted by the ERG to reflect specific patient ages; the results indicated that the combined effect of increasing mortality and reducing effectiveness could alter the estimated ICER substantially. For example, considering patient cohorts with ages at induction of 30, 60 and 90 years generated ICERs of £9,790, £18,055 and £43,306, respectively.

As a consequence of the immaturity of the PRIMA trial data [42] (RTX was not given to patients beyond 2 years), the manufacturer was obliged to model the long-term effect of RTX, making an assumption of the average period for which treatment continued to provide additional clinical effect. In the base case, the manufacturer considered a period of 6 years, but without reference to any supporting evidence. Three alternatives were put forward by the ERG: 26 months (the period over which data were available, i.e. 800 days), 4 years (equivalent to the maximum time over which any patients had been observed within the PRIMA trial [42]), and 40 years (equivalent to a lifetime). These time periods generated ICERs per QALY gained of £32,230, £21,151 and £8,966, respectively.

However, when considering these results it is important to note a major assumption used in the manufacturer's model, namely that, at the end of the assumed 'effective period', the hazard governing the PFS projective model reverts to the hazard of the comparator arm, but from a higher absolute survival level. This approach means that the survival curves of the intervention and comparator arms will never converge within a finite time, ensuring that the RTX arm continues to accumulate survival gains long after the supposed limit of clinical effectiveness. This assumption precludes an alternative process observed in some clinical trials, whereby clinical gains begin to decay at the end of active treatment, until the intervention and comparator survival curves converge to the same long-term trajectory after a few months or years. In simple terms, the difference can be characterized as follows:

- The manufacturer's approach assumes that the clock governing the disease process is 'turned back' for patients on RTX maintenance therapy by several years, so that it never catches up with that experienced by untreated patients.
- The alternative process assumes that the disease clock is slowed down by RTX for several years, but accelerates when the effect wears off, and eventually catches up with the disease experienced by the untreated patients.

The difference in health gain between these two modelling approaches was difficult to estimate using the manufacturer's model. However, in the manufacturer's base case, more than 72 % of the estimated PFS gain arises after the 4-year point; there is, therefore, considerable scope for a reduction in incremental outcomes if the initial PFS advantage is lost progressively. The estimated ICER could increase by as much as three times, depending on the time over which the gap between PFS curves disappears.

3.2.2 ERG Conclusions on the Submitted Cost-Effectiveness Evidence

Overall, the ERG considered that direct use of the model results was too heavily affected by extensive and often unquantifiable uncertainty around the central claims of clinical benefit to be useful for decision making, i.e. the clinical data available made it impossible to compare the cost effectiveness of first-line RTX maintenance therapy with observation in patients with fNHL with any confidence. In addition, several important model assumptions (the age of patients, the assumed duration of benefit from maintenance RTX therapy) indicated much greater uncertainty in the manufacturer's model cost-effectiveness results than was suggested in the MS.

3.3 Conclusions of the ERG Report

The ERG concluded that the results of the PRIMA trial [42] appeared promising; however, the early closure of the trial combined with limited follow-up resulted in an immature dataset. Too few patients had experienced events to draw any firm conclusions on clinical and cost effectiveness from the data available.

4 Key Methodological Issues

Clearly in this case, the data from the key clinical trial were immature and there was a lack of unequivocal data available to the manufacturer to support claims of PFS or OS benefit. Data from a longer follow-up would obviate the need for assumptions to be made around key variables. The ERG also considered that in the case of diseases such as fNHL that are of lengthy duration with complex treatment pathways, patient-level simulations rather than the Markov approach would better represent patient experience.

5 NICE Guidance

In June 2011, following three meetings, the AC recommended first-line RTX maintenance therapy as an option

for the treatment of people with fNHL whose disease has responded to first-line induction therapy with R-CTX.

5.1 Consideration of Clinical- and Cost-Effectiveness Issues

5.1.1 First AC meeting

5.1.1.1 Immature Trial Data The ERG's concerns regarding the immature results of the PRIMA trial [42] were discussed. The AC noted that the time to progression could not be estimated for patients in the RTX arm as too few patients had experienced disease progression at the time of data analysis. In addition, any OS benefit could not be calculated because of the small number of deaths. The AC also noted the early closure of the trial and considered evidence presented by the ERG that suggested that trials stopped earlier than planned often overestimate clinical benefit. However, following advice from clinical experts, the AC was satisfied that the PFS reported for patients in the RTX arm of the PRIMA trial [42] reflected observations from clinical practice. The AC concluded that the available evidence showed that first-line maintenance treatment with RTX significantly improves PFS compared with observation but that the size of any OS benefit could not be determined.

5.1.1.2 Appropriateness of the Economic Model The AC agreed that the manufacturer's model was acceptable to address the decision problem.

5.1.1.3 Duration of Clinical Benefit of Treatment The manufacturer's base-case assumption of 6 years of clinical benefit following RTX maintenance treatment was examined. Clinical experts informed the AC that the actual benefit is more likely to be around 3–4 years. The AC also acknowledged that when the ERG estimated the clinical benefit to last 3–4 years, the size of the ICER increased.

5.1.1.4 Conversion of PFS to OS The AC discussed the manufacturer's assumption of a 96.6 % conversion of PFS to OS and heard from the clinical specialists that the specific conversion factor cannot be verified from the literature or clinical experience. The clinical specialists considered a conversion closer to 70 % to be the most plausible. The AC acknowledged that the assumption of a conversion factor of 70 % would increase the base-case ICER. The AC was of the opinion that data from patient registries should have been used to validate the conversion factor assumed for the base-case estimate, and to confirm the degree to which RTX maintenance treatment might prolong life. The AC concluded that the manufacturer's base case underestimated the true ICER.

5.1.1.5 Patient Age The AC noted the mean age (56 years) of the patients in the PRIMA trial [42]. The

clinical specialists stated that the mean age of patients at the start of treatment in the UK is between 60 and 65 years. The AC acknowledged that people in clinical trials tend to be younger and fitter than those in clinical practice and noted from the ERG's suggested modification that the manufacturer's base-case ICER increased when the mean age at first treatment was assumed to be 60 and 65 years, respectively.

5.1.1.6 Utility Values The manufacturer's model included utility values of 0.88 and 0.79 for PFS1 and PFS2 health states, respectively. The AC noted that the ERG considered there should be no difference in the utility values used in the model to describe PFS1 and PFS2 as both groups of patients are in 'remission/full response'. When corrected by the ERG, the revision reduced the QALY gain in PRIMA [42] PFS by more than 10 % and therefore increased the ICER by approximately 11 %.

The AC agreed that the ICERs presented by the manufacturer were largely driven by gains in OS and considered that the model underestimated the utility associated with delaying CTX treatment and that if it was included, it would likely decrease the size of the ICER.

5.1.2 Summary and Outcome of the First AC Meeting

The AC agreed that the key drivers of cost effectiveness were (1) the assumptions concerning the duration of clinical benefit of maintenance RTX therapy, (2) the conversion of PFS to OS and (3) the underestimation in the economic model of the utility associated with delaying CTX treatment following the first relapse.

The AC concluded that it was difficult to establish the most plausible ICER for RTX maintenance treatment because SAs to capture preferred assumptions were not available; however, the ICER may be within the range that is consistent with an appropriate use of NHS resources.

The first ACD [52] recommended the use of first-line maintenance RTX treatment for patients with fNHL who have responded to induction R-CTX therapy provided that the manufacturer addressed the uncertainties within a revised cost-effectiveness analysis:

- for people who are aged 60 to 65 years at the start of treatment, and
- in which the duration of clinical benefit from [RTX] maintenance treatment is 3 to 4 years, and
- in which the extent that [PFS] translates into [OS] is informed by the best available evidence, which could include patient registries and prospective observational data. Analyses should be presented for a range of plausible values, from 50 % to 100 %, and

- in which utility gains associated with delaying the need for [CTX] after relapse are included [52].

5.1.3 Second AC Meeting

The AC concluded that the ICERs presented in the manufacturer's revised submission were associated with a great deal of uncertainty. Further clarification was sought from the manufacturer as follows:

1. Exploration of the duration of benefit at different intervals (28 months, 36 months or 48 months)
2. Additional analyses that assume a conversion factor of 70 %, 80 % or 90 %
3. The mean age of patients at the start of treatment to be 62.5 years [52]

At the end of the second meeting, the AC was minded not to recommend RTX maintenance treatment for this indication as a cost-effective use of NHS resources.

5.1.4 Third AC Meeting

The AC considered the manufacturer's SAs that assumed a duration of treatment effect of 28 months, 36 months and 48 months and noted that the ICERs ranged from £17,300 to £27,400 per QALY gained. It was noted that the estimates were lower than the ERG scenarios; however, the manufacturer and the ERG had used different modelling approaches. The AC was satisfied that the manufacturer's SAs presented the most plausible range of estimates for the treatment effect in line with clinical opinion and the available data. The AC was also satisfied that the manufacturer's SAs, which assumed conversion rates of 70 %, 80 % and 90 %, provided a plausible range of conversion rate estimates.

In its final deliberations, the AC noted that the manufacturer's ICERS for RTX maintenance therapy compared with observation were less than £30,000 per QALY gained for most scenarios and the ERG ICERs ranged from £24,600 to £35,000 per QALY gained (depending on the conversion rate of PFS to OS). The AC also noted that the model did not include the utility associated with delaying CTX, and that if it were included it would decrease the ICER to an estimate that would be considered as a cost-effective use of NHS resources. The AC therefore recommended the use of RTX in this indication.

6 Conclusion

The main problem in this appraisal was the uncertainty of the clinical and cost effectiveness of RTX as a maintenance

treatment for patients with fNHL whose disease had responded to induction treatment with R-CTX. The uncertainty was a direct result of the immature data presented as evidence of clinical effectiveness. The reported data appeared promising, but the lack of events mitigated against drawing any firm conclusions for either clinical or cost effectiveness. The early closure of the trial was of further concern given that evidence indicates that trials stopped early tend to overestimate treatment effects [46].

At the time that the ERG report was written, the median PFS for the RTX arm of the trial was not estimable; however, PFS was the key driver of the economic model. In the view of the ERG, the manufacturer was unable to predict the duration of the clinical effect of RTX maintenance therapy on PFS with any plausible certainty. Additionally, the manufacturer was unable to justify any assumptions regarding a direct survival benefit from increased PFS as the PFS to OS gain ratio in this disease area is unknown.

The ERG maintained the position throughout the process that the trial data were too immature to derive any firm conclusions regarding the clinical and cost effectiveness of RTX as a maintenance treatment and considered that any re-manipulation of the limited data available would not address any of the existing uncertainties. Only a fully mature dataset would adequately address these. As noted by the EMA [53] in its evaluation of the data from the PRIMA trial [42], a median follow-up of about 2 years is relatively short for the assessment of the long-term benefit of RTX maintenance treatment. The ERG awaits the publication of more mature data from the PRIMA trial [42].

Acknowledgments This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment Programme (project number 08/218) and is published as part of a compendium of ERG articles in *Health Technology Assessment* [54]. See the Health Technology Assessment Programme website for further project information: <http://www.hta.ac.uk>. This summary of the ERG report was compiled after the AC's consideration of the evidence and has not been externally peer reviewed by *PharmacoEconomics*. The views and opinions expressed herein are those of the authors and do not necessarily reflect those of the NICE or the Department of Health. The authors have no competing interests.

Author Contributions Janette Greenhalgh led the project, drafted the Clinical Effectiveness section and supervised the production of the manuscript. Adrian Bagust was responsible for the critical appraisal of the economic model. Angela Boland and Sophie Beale were responsible for the critical appraisal of the economic evidence. Michaela Blundell and James Oyee were responsible for the critical appraisal of the clinical statistical approach. Yenal Dundar cross-checked the manufacturer's search strategies. Juliet Hockenhuil was responsible for the summary and critical appraisal of the clinical evidence. Chris Proudlove was involved in the critical appraisal of the MS. Patrick Chu was responsible for the critical appraisal of the clinical sections of the MS. All authors read and commented on draft versions of the manuscript. Janette Greenhalgh is the guarantor for the overall content.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited. The exclusive right to any commercial use of the article is with Springer.

References

- National Institute for Health and Clinical Excellence. Guide to the single technology (STA) process. London: NICE; 2006. <http://www.nice.org.uk/page.aspx?o=STAProcessguide>. Accessed 1 Sep 2010.
- Sculpher M. Single technology appraisal at the UK National Institute for Health and Clinical Excellence: a source of evidence and analysis for decision making internationally. *Pharmacoeconomics*. 2010;28(5):347–9.
- Rodgers M, Griffin S, Paulden M, et al. Alitretinoin for severe chronic hand eczema: a NICE single technology appraisal. *Pharmacoeconomics*. 2010;28(5):351–62.
- Bagust A, Greenhalgh J, Boland A, et al. Cetuximab for recurrent and/or metastatic squamous cell carcinoma of the head and neck: a NICE single technology appraisal. *Pharmacoeconomics*. 2010; 28(6):439–48.
- Stevenson M, Pandor A. Febuxostat for the management of hyperuricemia in patients with gout: a NICE single technology appraisal. *Pharmacoeconomics*. 2011;29(2):133–40.
- Scotland G, Waugh N, Royle P, et al. Denosumab for the prevention of osteoporotic fractures in post-menopausal women: a NICE single technology appraisal. *Pharmacoeconomics*. 2011;29(11): 951–61.
- Dickson R, Bagust A, Boland A, et al. Erlotinib monotherapy for the maintenance treatment of non-small cell lung cancer after previous platinum-containing chemotherapy: a NICE single technology appraisal. *Pharmacoeconomics*. 2011;29(12):1051–62.
- McKenna C, Maund E, Sarowar M, et al. Dronedarone for the treatment of atrial fibrillation: a NICE single technology appraisal. *Pharmacoeconomics*. 2012;30(1):35–46.
- Holmes M, Carroll C, Papaioannou D. Dabigatran etexilate for the prevention of venous thromboembolism in patients undergoing elective hip and knee surgery: a NICE single technology appraisal. *Pharmacoeconomics*. 2012;30(2):137–46.
- Yang H, Craig D, Epstein D, et al. Golimumab for the treatment of psoriatic arthritis: a NICE single technology appraisal. *Pharmacoeconomics*. 2012;30(4):257–70.
- Boyers D, Jia X, Jenkinson D, et al. Eltrombopag for the treatment of chronic immune or idiopathic thrombocytopenic purpura: a NICE single technology appraisal. *Pharmacoeconomics*. 2012;30(6):483–95.
- Burch J, Griffin S, McKenna C, Walker S, Paton J, Wright K, et al. Omalizumab for severe persistent asthma in children aged 6–11 years: a NICE single technology appraisal. *Pharmacoeconomics*. 2012;30(11):991–1004.
- Whyte S, Pandor A, Stevenson M. Bevacizumab for metastatic colorectal cancer: a NICE single technology appraisal. *Pharmacoeconomics*. 2012;30(12):1119–32.
- Kilonzo M, Hislop J, Elders A, Fraser C, Bissett D, McClinton S, et al. Pazopanib for the first-line treatment of patients with advanced and/or metastatic renal cell carcinoma: a NICE single technology appraisal. *Pharmacoeconomics*. 2013;31(1):15–24.
- Craig D, Rice S, Paton F, et al. Retigabine for the adjunctive treatment of adults with partial onset seizures in epilepsy with and without secondary generalisation: a NICE single technology appraisal. *Pharmacoeconomics*. 2013;31(2):101–10.
- Spackman E, Rice S, Norman G, Suh D-C, Eastwood A, Palmer S. Trastuzumab for the treatment of HER2 positive metastatic gastric cancer: a NICE single technology appraisal. *Pharmacoeconomics*. 2013;31(3):185–94.
- Rafia R, Simpson E, Stevenson M, Papaioannou D. Trabectedin for the treatment of advanced metastatic soft tissue sarcoma: a NICE single technology appraisal. *Pharmacoeconomics*. 2013. doi:10.1007/s40273-013-0044-7.
- Simpson EL, Fitzgerald P, Evans P, et al. Bivalirudin for the treatment of ST-segment elevation myocardial infarction: a NICE single technology appraisal. *Pharmacoeconomics*. 2013;31(4):269–75.
- Armstrong N, Manuela J, van Asselt T, et al. Golimumab for the treatment of ankylosing spondylitis: a NICE single technology appraisal. *Pharmacoeconomics*. 2013. doi:10.1007/s40273-013-0049-2.
- Tosh J, Archer R, Davis S, et al. Golimumab for the treatment of rheumatoid arthritis after the failure of previous disease-modifying anti-rheumatic drugs: a NICE single technology appraisal. *Pharmacoeconomics* (in press).
- Kearns B, Lloyd-Jones M, Stevenson M, Littlewood C. Cabazitaxel for the second-line treatment of metastatic hormone-refractory prostate cancer: a NICE single technology appraisal. *Pharmacoeconomics*. 2013. doi:10.1007/s40273-013-0050-9.
- Faria R, Spackman E, Burch J, Corbacho B, Todd D, Pepper C, et al. Dabigatran for the prevention of stroke and systemic embolism in atrial fibrillation: a NICE single technology appraisal. *Pharmacoeconomics* (in press).
- National Institute for Health and Clinical Excellence. Lymphoma (follicular non-Hodgkin's)—rituximab: final appraisal determination. London: NICE; 2011. <http://guidance.nice.org.uk/TA/Wave19/59>. Accessed 1 Sep 2011.
- Cancer Research UK. Non-Hodgkin lymphoma (NHL) statistics (UK 2010). London: Cancer Research UK; 2010. <http://info.cancerresearchuk.org/cancerstats/types/nhl/index.htm?script=true>. Accessed 1 Sep 2010.
- National Cancer Institute, U.S. National Institutes of Health. Surveillance Epidemiology and End Results: SEER stat fact sheets—non-Hodgkin lymphoma. Bethesda: National Cancer Institute; 2012. <http://seer.cancer.gov/statfacts/html/nhl.html>. Accessed 1 May 2012.
- Australian Institute of Health and Welfare, Australasian Association of Cancer Registries. Cancer in Australia: an overview, 2008 [cancer series no.46]. Canberra: AIHW; 2008. <http://www.aihw.gov.au/publication-detail/?id=6442468196>. Accessed 1 May 2012.
- Non-Hodgkin's Lymphoma Cyberfamily. NHL statistics. <http://www.nhlcyberfamily.org/statistics.htm#england>. 2012. Accessed 1 May 2012.
- A predictive model for aggressive non-Hodgkin's lymphoma. The International non-Hodgkin's Lymphoma Prognosis Factors Project. *N Engl J Med*. 1993;329(14):987–94.
- National Institute for Health and Clinical Excellence. Follicular non-Hodgkin's lymphoma—rituximab: final scope. London: NICE; 2010. <http://guidance.nice.org.uk/TA/Wave19/59/Scope/pdf/English>. Accessed 25 Aug 2010.
- Shipp MA, Mauch PM, Harris NL, editors. Non-Hodgkin's lymphomas. 5th ed. Philadelphia: Lippincott-Raven; 1997.
- Horning SJ, Rosenberg S. Natural history of initially untreated low-grade non-Hodgkin's lymphoma. *N Engl J Med*. 1984;311:1471–5.
- Lister T. The management of follicular lymphoma. *Ann Oncol*. 1991;2(Suppl. 2):131–5.
- Dillman RO, Chico S. Improved survival of lymphoma patients after introduction of rituximab. *Blood*. 2005;106: abstract 4650.
- Fischer RI, LeBlanc M, Press OW, Maloney DG, Unger JM, Miller TP. New treatment options have changed the survival of patients with follicular lymphoma. *J Clin Oncol*. 2005;23(33): 8447–52.

35. Liu Q, Fayad L, Hagemeister FB, Alma Rodriguez M, Younes A, Pro B, et al. Stage IV indolent lymphoma: 25 years of treatment progress. *Blood*. 2003;102(398a): abstract 1446.
36. Swenson WT, Wooldridge JE, Lynch CF, Forman-Hoffman VL, Chrischilles E, Link BK. Improved survival of follicular lymphoma patients in the United States. *J Clin Oncol*. 2005;23: 5019–26.
37. Gallagher CJ, Gregory WM, Jones AE, Stansfeld AG, Richards MA, Dhaliwal HS, et al. Follicular lymphoma: prognostic factors for response and survival. *J Clin Oncol*. 1986;4:1470–80.
38. Wahl R, Kaminski M, Zelenetz A, Vose J, Press O, Goldsmith S, et al. Each subsequent therapy results in diminishing response rate and duration of response in low grade or transformed low grade non-Hodgkin's lymphoma. *Proc Am Soc Clin Oncol*. 2001; 20: abstract 1165.
39. van Oers M, Klasa R, Marcus RE, Wolf M, Kimby E, Gascoyne RD, et al. Rituximab maintenance improves clinical outcome of relapsed/resistant follicular non-Hodgkin lymphoma in patients both with and without rituximab during induction: results of a prospective randomized phase 3 intergroup trial. *Blood*. 2006; 108(10):3295–301.
40. van Oers MHJ, van Glabbeke M, Giurgea L, Klasa R, Marcus RE, Wolf M, et al. Rituximab maintenance treatment of relapsed/resistant follicular non-Hodgkin's lymphoma: long-term outcome of the EORTC 20981 phase III randomized intergroup study. *J Clin Oncol*. 2010;28(17):2853–8.
41. European Medicines Agency. Mabthera (rituximab). London: EMA; 2012. http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/000165/human_med_000897.jsp&mid=WC0b01ac058001d124. Accessed 1 May 2012.
42. Roche Ltd. Rituximab: clinical study report (M018264)-No.10347952010.
43. Salles G, Seymour JF, Offner F, Lopez-Guillermo A, Belada D, Xerri L, et al. Rituximab maintenance for 2 years in patients with high tumour burden follicular lymphoma responding to rituximab plus chemotherapy (PRIMA): a phase 3, randomised controlled trial. *Lancet*. 2011;377(9759):42–51.
44. Functional Assessment of Chronic Illness Therapy. FACIT. Elmhurst: FACIT; 2010. <http://www.facit.org/FACITOrg/Questionnaires>. Accessed 1 Sep 2010.
45. European Organisation for Research and Treatment of Cancer. EORTC QLQ-C30. Brussels: EORTC; 2010. <http://groups.eortc.be/qol/eortc-qlq-c30>. Accessed 1 Sep 2010.
46. Bassler D, Briel M, Montori VM, Lane M, Glasziou P, Zhou Q, et al. Stopping randomized trials early for benefit and estimation of treatment effects: systematic review and meta-regression analysis. *JAMA*. 2010;303(12):1180–7.
47. National Institute for Health and Clinical Excellence. Rituximab for the treatment of relapsed or refractory stage III or IV follicular non-Hodgkin's lymphoma (review of technology appraisal guidance 37):TA137. London: NICE; 2008. <http://guidance.nice.org.uk/TA137/Guidance/pdf/English>. Accessed 25 Aug 2010.
48. Wild D. Utility values in follicular lymphoma - Oxford Outcomes [unpublished report by Oxford Outcomes prepared for Roche UK]. Oxford: Oxford Outcomes; 2005.
49. Department of Health. NHS reference costs 2008-2009. London: Department of Health; 2010. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_111591. Accessed 25 Aug 2010.
50. Royal Pharmaceutical Society of Great Britain. British National Formulary: BNF 56-59. London: BMJ Publishing, RCPCH Publications, Royal Pharmaceutical Society of Great Britain; 2010. <http://www.bnf.org/bnf/>. Accessed 6 July 2010.
51. Nabhan C, Smith SM, Kahl BS. Maintenance rituximab in follicular non-Hodgkin lymphoma: facts and controversies. *Leukemia Lymphoma*. 2012;53(5):770–8.
52. National Institute for Health and Clinical Excellence. Follicular non-Hodgkin's lymphoma—rituximab: appraisal consultation document. London: NICE; 2010. <http://guidance.nice.org.uk/TA/Wave19/59/Consultation/DraftGuidance1>. Accessed 1 Oct 2011.
53. European Medicines Agency. Assessment report for MabThera (rituximab) [procedure no. EMEA/H/C/000165/II/0069]. London: EMA; 2010. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Assessment_Report_-_Variation/human/000165/WC500099489.pdf. Accessed 1 Oct 2011.
54. Greenhalgh J, Bagust A, Boland A, Blundell M, Oyee J, Dundar Y, et al. Rituximab for the first-line maintenance treatment of follicular non-Hodgkin's lymphoma. Southampton: NETSCC; 2012.
55. Liverpool Reviews and Implementation Group. Rituximab for the first-line maintenance treatment of patients with follicular non-Hodgkin's lymphoma: addendum [ERG critique of extra information submitted by Roche]. London: NICE; 2011. <http://www.nice.org.uk/nicemedia/live/11863/54540/54540.pdf>. Accessed 1 Oct 2011.