Article:
Tappenden, P., Simpson, E., Hamilton, J. et al. (2019) Ibrutinib for Treating Relapsed or Refractory Mantle Cell Lymphoma: An Evidence Review Group Perspective of a NICE Single Technology Appraisal. PharmacoEconomics, 37 (3). pp. 333-343. ISSN: 1170-7690

https://doi.org/10.1007/s40273-018-0713-7

Ibrutinib for treating relapsed or refractory mantle cell lymphoma: An Evidence Review Group perspective of a NICE Single Technology Appraisal

List of authors

Paul Tappenden, Reader in Health Economic Modelling, ScHARR, University of Sheffield, UK

Emma Simpson, Research Fellow, ScHARR, University of Sheffield, UK

Jean Hamilton, Research Fellow, ScHARR, University of Sheffield, UK

Daniel Pollard, Research Associate, ScHARR, University of Sheffield, UK

Mark Clowes, Information Specialist, ScHARR, University of Sheffield, UK

Eva Kaltenthaler, Professor of Health Technology Assessment, ScHARR, University of Sheffield, UK

David Meiklejohn, Consultant Haematologist, Department of Haematology, Ninewells Hospital, UK

Nick Morley, Consultant Haematologist, Sheffield Teaching Hospitals NHS Foundation Trust, UK

Corresponding author

Dr Paul Tappenden, Reader in Health Economic Modelling, ScHARR, University of Sheffield,
Regent Court, 30 Regent Street, Sheffield, S1 4DA, England.

Tel: +44 114 2220855

Fax: +44 114 2724095

Email: p.tappenden@sheffield.ac.uk

Running header: ERG review of ibrutinib for mantle cell carcinoma

ABSTRACT

As part of its Single Technology Appraisal process, the UK National Institute for Health and Care Excellence (NICE) invited the manufacturer of ibrutinib (Janssen) to submit evidence on the clinical effectiveness and cost-effectiveness of ibrutinib for the treatment of relapsed or refractory (R/R) mantle cell lymphoma (MCL). The School of Health and Related Research Technology Assessment Group at the University of Sheffield was commissioned to act as the independent Evidence Review Group (ERG). The ERG produced a critical review of the evidence contained within the company's submission to NICE. The clinical effectiveness evidence for ibrutinib included one randomised controlled trial (RCT) comparing ibrutinib versus temsirolimus (TEM) and two single-arm studies. The company's indirect comparison of ibrutinib versus rituximab plus chemotherapy (R-chemo) produced a hazard ratio (HR) for progression-free survival (PFS) of 0.28. The ERG's random effects network meta-analysis (NMA) indicated that the treatment effect on PFS is highly uncertain (HR=0.27; 95% credible interval [CrI] 0.06, 1.26). The company's Markov model assessed the cost-effectiveness of ibrutinib versus R-chemo for the treatment of R/R MCL from the perspective of the National Health Service (NHS) and Personal Social Services (PSS) over a lifetime horizon. Based on a re-run of the company's model by the ERG, the incremental cost-effectiveness ratio (ICER) for ibrutinib versus R-chemo (including the company's original Patient Access Scheme [PAS]) was expected to be £76,014 per quality-adjusted life year (QALY) gained. The ERG had several concerns regarding the company's model structure and the evidence used to inform its parameters. The ERG's preferred analysis, which used the ERG's NMA and the observed Kaplan-Meier curve for time to ibrutinib discontinuation and which excluded long-term disutilities for R-chemo, produced ICERs of £63,340 per QALY gained for the overall R/R MCL population, and £44,711 per QALY gained for patients with one prior treatment. Following an updated PAS and consideration of evidence from a later data-cut of the RAY trial, the appraisal committee concluded that the most plausible ICER for the one prior treatment subgroup is likely to be lower than the company's estimate of £49,848 per QALY gained. The company's ICER for the overall R/R MCL population was higher, at £62,650 per QALY gained. The committee recommended ibrutinib as an option for treating R/R MCL in adults, only if they have had only one previous line of therapy and the company provides ibrutinib with the discount agreed in the commercial access agreement with NHS England.

KEY POINTS FOR DECISION-MAKERS

- The clinical evidence for ibrutinib was comprised of one randomised controlled trial of ibrutinib versus temsirolimus (RAY [MCL3001]) and two single-arm studies (PCYC1104 and SPARK [MCL2001]).
- The ERG's main concerns related to uncertainty surrounding the relative benefits of ibrutinib versus treatments currently used in UK clinical practice. The ERG's exploratory random effects network meta-analysis suggested a hazard ratio for PFS for ibrutinib versus rituximab plus chemotherapy of 0.27 (95% credible interval 0.06 to 1.23).
- The ERG's preferred analysis of the company's model (based on the original Patient Access Scheme [PAS]) resulted in an incremental cost-effectiveness ratio (ICER) of £63,340 per quality-adjusted life year (QALY) gained for the relapsed/refractory (R/R) mantle cell lymphoma (MCL) population and £44,711 per QALY gained for the one prior therapy subgroup.
- The committee concluded that the most plausible ICER for the one prior therapy subgroup is likely to be lower than £49,848 per QALY gained. Ibrutinib was recommended as an option for treating R/R MCL in adults, only if they have had only one previous line of therapy and the company provides ibrutinib with the discount agreed in the commercial access agreement with NHS England.

1. INTRODUCTION

Health technologies must be shown to be clinically effective and to represent a cost-effective use of resources to be recommended for use within the National Health Service (NHS) in England. The National Institute for Health and Care Excellence (NICE) is an independent organisation responsible for providing national guidance on promoting good health, and preventing and treating ill health, in priority areas with significant impact. The NICE Single Technology Appraisal (STA) process usually covers new technologies soon after they have received UK marketing authorisation and is specifically designed for the appraisal of a single health technology within a single indication [1]. Within the STA process, the manufacturer of a technology provides NICE with a written submission containing relevant clinical effectiveness evidence alongside a health economic model that summarises the company's estimates of the cost-effectiveness of the technology. The Evidence Review Group (ERG), an external academic organisation which is independent of NICE, reviews the submission with advice from clinical specialists and produces an ERG report. The NICE appraisal committee (AC) considers the company's submission (CS), the ERG report, and testimony from experts and other stakeholders and formulates preliminary guidance - the appraisal consultation document (ACD) - which indicates the initial decision of the AC regarding the recommendation (or not) of the intervention. Stakeholders are subsequently invited to comment on the submitted evidence and the ACD, after which an ACD may be produced or a final appraisal determination (FAD) issued, which is open to appeal. An ACD is not produced when the intervention is recommended without restriction; in such instances, a FAD is produced directly. This paper presents a summary of the ERG report produced by the School of Health and Related Research Technology Assessment Group at the University of Sheffield [2] and the NICE FAD [3] for the STA of ibrutinib for the treatment of patients with relapsed or refractory (R/R) mantle cell lymphoma (MCL). It also covers the subsequent development of the NICE guidance for the use of this drug in England [4]. Full details of all relevant appraisal documents can be found on the NICE website [5].

2. THE DECISION PROBLEM

MCL is a B-cell malignancy with unique biological, pathological and clinical features, which comprises approximately 3-10% of all non-Hodgkin lymphomas (NHLs). MCL typically arises in older adults, with a median age of presentation of between 60 and 65 years of age. Approximately 75% of patients with MCL are male. The disease is rare: the company estimated that there would be 356 patients with R/R MCL in England, Wales and Northern Ireland in 2017 [6]. MCL is characterised by an aggressive clinical disease course, but features a pattern of resistant and relapsing disease, thereby rendering it incurable with standard therapy. The prognosis for patients with MCL is very poor compared with other forms of NHL; after excluding patients for whom autologous stem cell transplantation (ASCT) is a treatment option, median survival for patients with MCL following first relapse is typically reported to be approximately 1-2 years [7], but may be lower.

2.1 Current treatment

There is no standard of care for patients with R/R MCL and the disease remains very difficult to manage. The British Committee for Standards in Haematology (BCSH) guidelines [7] note that the choice of therapy at relapse will be determined by patient age, performance status, bone marrow reserve, initial therapy and history of infections. For patients who have not received transplantation as first-line therapy, but are sufficiently fit for such therapy following relapse, ASCT may be considered as a clinical option [7]. However, this is not a common scenario as most patients who are suitable for an autograft will have received it first-line when it is best tolerated and likely to be most effective. For older and/or less fit patients, a range of systemic chemotherapy regimens may be considered, provided the patient is sufficiently fit to receive them. Treatment typically involves chemotherapy with or without rituximab and may include regimens such as: R-CHOP (rituximab, cyclophosphamide, doxorubicin and vincristine); R-bendamustine (rituximab and bendamustine); FCR (rituximab, fludarabine and cyclophosphamide); R-CVP (rituximab, cyclophosphamide, vincristine and prednisolone), or R-chlorambucil (rituximab and chlorambucil). None of these therapies are specifically licensed for the treatment of R/R MCL. Guidelines published by the European Society for Medical Oncology (ESMO) discourage the use of R-CVP and FCR due to inferior response rates and long-lasting myelosuppression [8]. The BCSH guidelines for MCL report that several other regimens have been shown to have activity in R/R MCL, including: (i) bortezomib; (ii) bortezomib-gemcitabine; (iii) bortezomib-rituximab; (iv) temsirolimus (TEM); (v) rituximab, fludarabine, cyclophosphamide and mitoxantrone (R-FCM); (vi) fludarabine; (vii) fludarabine-cyclophosphamide; (viii) fludarabine-cyclophosphamide+/-rituximab; (ix) cladribine; (x) gemcitabine-dexamethasone; (xi) gemcitabine-dexamethasone-cisplatin; (xii) lenalidomide; (xiii) thalidomide-rituximab, and (xiv) prednisone, etoposide, procarbazine, and cyclophosphamide (PEP-C)/thalidomide/rituximab [7]. Where patients have received one previous line of treatment, a different regimen would typically be used following relapse. Whilst ESMO recommends targeted therapies such as TEM, bortezomib and lenalidomide [8], TEM is not currently used in England, bortezomib is recommended only in untreated MCL and lenalidomide is available only via a compassionate use programme [6].

In February 2016, NICE issued a final scope to appraise the clinical effectiveness and cost-effectiveness of ibrutinib for the treatment of R/R MCL [9].

3. INDEPENDENT ERG REVIEW

The company (Janssen) provided a submission to NICE on the clinical effectiveness and cost-effectiveness of ibrutinib for the treatment of R/R MCL [6]. This submission was critically appraised by the ERG. Subsequently, the ERG identified areas requiring clarification, for which the company provided additional evidence prior to completion of the ERG report [10].

3.1 Clinical evidence submitted by the company

Clinical effectiveness data were taken from three studies, comprising one randomised controlled trial (RCT) (RAY [MCL3001]) and two single-arm studies (PCYC1104 and SPARK [MCL2001]) [11-13]. One additional study of TEM versus physicians' choice of therapy (the OPTIMAL study [14]) was included in an indirect comparison of ibrutinib versus single-agent chemotherapy.

At the time of the CS, median OS had not been reached in the ibrutinib arm of the RAY study (n=139) or in the SPARK study (n=120). In Study PCYC1104 (n=111), median OS was 22.5 months.

Median PFS for ibrutinib-treated patients was 14.6 months in the RAY study, 13.0 months in the final analysis of Study PCYC1104, and 10.5 months in the SPARK study. In the RAY study, median PFS within the TEM arm was 6.2 months; this was significantly worse than for the ibrutinib arm (hazard ratio [HR]=0.43, 95% confidence interval [CI] 0.32 to 0.58; $p<0.0001$). Overall response rates (ORR) assessed by independent review committee (IRC) were similar for ibrutinib-treated patients across studies (71.9% in RAY, 69% in Study PCYC1104 and 62.7% in SPARK). In RAY, there was a significant advantage in overall response rate (ORR) for ibrutinib over TEM (ORR=40.4%, odds ratio [OR]=3.98, 95% CI 2.38 to 6.65; difference in ORR $p<0.0001$).

Health-related quality of life (HRQoL) was measured by Functional Assessment of Cancer Therapy - Lymphoma (FACT-Lym) in the RAY (MCL3001) and SPARK (MCL2001) studies [15]. In the RAY study, 61.9% of ibrutinib-treated patients reported a clinically meaningful improvement; significantly fewer TEM-treated patients reported clinically meaningful improvement (35.5%, $p<0.0001$).

Across the three studies of ibrutinib (RAY, SPARK and PCYC1104), the most common adverse events (AEs) for ibrutinib (≥20% of patients) were: diarrhoea; cough; fatigue; thrombocytopenia; neutropenia; peripheral oedema; nausea; muscle spasms, and pyrexia.

The company's indirect comparison of ibrutinib versus single-agent chemotherapy suggested that ibrutinib is associated with a slower rate of disease progression compared with single-agent chemotherapy (HR=0.19, 95% CI 0.10, 0.36) and a survival benefit (HR=0.61, 95% CI 0.34, 1.10), although the result for OS is inconclusive as it did not reach statistical significance at the 95% level. Rituximab is used in routine clinical practice in England, therefore to account for the differential effectiveness of using rituximab alongside chemotherapy, the company performed an additional adjustment to the HR for PFS. The adjusted HR for PFS for ibrutinib versus R-chemo was estimated to be 0.28.

3.2 Critique of clinical effectiveness evidence and interpretation

The ERG believed that all relevant studies had been included in the CS [6]. The studies presented were relevant to the population, intervention and outcomes of the decision problem. The CS did not identify any RCTs which included head-to-head comparisons of ibrutinib versus any of the comparators listed in the final NICE scope [9]. TEM, the comparator in the RAY trial, is not used in UK practice.

The populations of the three included ibrutinib studies reflect the demographic characteristics of the R/R MCL population that would be eligible for treatment using ibrutinib in England. However, in practice, patients may have more comorbidities than those included in the studies. The studies were international, with a small proportion of patients from the UK, thus there may be differences between the treatment pathways of the patients enrolled into the ibrutinib studies and those seen in current practice in England.

One of the included ibrutinib studies was an RCT (RAY), whereas the other two studies (SPARK and PCYC1104) adopted a single-arm design. All three included ibrutinib studies were open-label and therefore were subject to potential bias. However, all studies addressed the issue of measurement bias for the primary outcome by having an assessment of the primary outcome by IRC. All three studies were sufficiently large to be adequately powered for their primary endpoint of PFS (RAY) or ORR (PCYC1104 and SPARK). OS was not adequately powered, and may have been influenced by the differential use of subsequent therapies between treatment groups. The TEM arm in the RAY study had better outcomes than the TEM arm in the OPTIMAL study. There is uncertainty regarding how much of this difference is due to TEM treatment, differences in populations between trials and routine practice, and the use of other therapies, or random chance.

The company adopted a two-stage approach to estimate treatment effects for ibrutinib versus R-chemo. The ERG considers that a single-stage random effects NMA would provide a better representation of the uncertainty in the resulting treatment comparisons. Based on additional analyses undertaken by the ERG, ibrutinib is associated with a slower rate of disease progression, compared to R-chemo, but with considerable uncertainty (random effects HR=0.27, 95% credible interval [CrI] 0.06 to 1.26). The estimated median HRs for OS for ibrutinib versus R-chemo ranged from 0.98 to 1.96, depending on the data source used for the rituximab arm of the network (an RCT reported by Forstpointner *et al* [16], audit data from the Haematological Malignancy Research Network [HMRN] [17], or both). Due to concerns regarding the evidence used to inform the indirect comparisons, the ERG advised that the results of the indirect comparison should be interpreted with caution.

3.3 Cost-effectiveness evidence submitted by the company

The company submitted a *de novo* health economic model which assessed the cost-effectiveness of ibrutinib versus R-CHOP (rituximab, cyclophosphamide, vincristine and prednisolone) for the treatment of patients with R/R MCL over a 15-year (lifetime) horizon from the perspective of the UK NHS and Personal Social Services (PSS). Costs and health outcomes were discounted at a rate of 3.5% per annum. Unit costs were valued at 2014/15 prices. Separate subgroup analyses were presented for patients who have received one prior line of therapy (LOT) and for patients who have received ≥2 prior LOTs. At the time of the appraisal, a Patient Access Scheme (PAS) in the form of a simple price discount had been agreed with the Department of Health. The value of this discount changed over the course of the appraisal. The level of the PAS is confidential; all cost-effectiveness results presented here are based on the original PAS, unless otherwise stated.

The company's base case model included three health states: (i) progression-free; (ii) post-progression, and; (iii) dead. The model also implicitly included a further separation between patients who are progression-free and on treatment and those who are progression-free after treatment discontinuation. The model adopted a 28-day cycle length. Within the ibrutinib group, health state transitions were modelled using parametric survivor functions fitted to data on pre-progression-mortality (exponential model), PFS (Weibull model), and post-progression survival (PPS, exponential model) from a pooled dataset of the RAY, SPARK and PCYC1104 studies [18]. Time to treatment discontinuation or death (TTD/D) was also modelled using a parametric survivor function (Weibull model), but did not impact on transitions between health states. The benefits of ibrutinib versus R-CHOP were modelled using the company's indirect comparison using the RAY trial [19] (ibrutinib versus TEM), the OPTIMAL trial [14] (TEM versus physician's choice of single-agent chemotherapy) and audit data from the HMRN [17] (R-chemo versus chemotherapy alone). This HR was applied to both the PFS and TTD/D curves for ibrutinib. The PPS curves for both groups were based on the pooled ibrutinib dataset [18]. Health utilities for the progression-free and post-progression states were derived from EQ-5D-5L data collected within the RAY and SPARK studies [19, 20]; the model also included a disutility associated with R-chemo toxicity which was based on clinical opinion. Health utilities were age-adjusted. The company's model included costs associated with drug acquisition, drug administration, follow-up, management of AEs, best supportive care (BSC) and death. Other resource use, including imaging, tests, biopsies, transfusions and hospitalisations, was estimated from a survey of NHS haematologists and oncologists [6]. Unit costs were taken from the Monthly Index of Medical Specialties (MIMs) [21], the Commercial Medicines Unit (CMU) electronic market information tool (eMit) [22] and NHS Reference Costs [23].

Based on a re-run of the probabilistic version of the company's base case model by the ERG, the incremental cost-effectiveness ratio (ICER) for ibrutinib versus R-CHOP was expected to be £76,014 per quality-adjusted life year (QALY) gained. The deterministic model produced similar results

(ICER=£75,317 per QALY gained). Assuming a willingness-to-pay (WTP) threshold of £50,000 per QALY gained, the company's base case model suggested that the probability that ibrutinib produces more net benefit than R-CHOP was approximately zero. Across all but one of the company's scenario analyses, the ICER for ibrutinib versus R-chemo was greater than £70,000 per QALY gained. The only exception to this related to an analysis in which the PPS rate for R-chemo was "calibrated" such that the overall modelled OS for R-CHOP was equal to the 8.4 month OS estimate reported within the HMRN audit [17] (ICER=£59,345 per QALY gained). However, it should be noted that this was an analysis in the one prior LOT subgroup rather than the overall population.

3.3.1 Critique of cost-effectiveness evidence and interpretation

The ERG critically appraised the company's economic analysis and double-programmed the deterministic version of the company's model. The main issues identified by the ERG are discussed below; the full critique can be found in the ERG report [2].

3.3.1.1 Modelling approach

The company's Markov model used three sets of transition probabilities derived from time-to-event data: (i) progression-free to dead (pre-progression mortality); (ii) progression-free to post-progression (calculated using PFS adjusted for pre-progression mortality), and; (iii) post-progression to dead (PPS). Whilst Markov models are commonly used in many areas of evaluation, the ERG had three main concerns with the company's approach: (a) the hazard of pre-progression mortality was assumed to be constant over time; (b) the use of PPS data includes only those patients who have progressed and thus may be subject to selection bias, and (c) the model included a structural assumption that PPS was exponentially distributed. Furthermore, because the same PPS curve was used for both treatment groups, this use of evidence within a Markov framework imposes a direct surrogate relationship between PFS gains and OS gains, whereby if the same pre-progression mortality curve is assumed for both groups, the incremental gain in PFS leads to an equivalent incremental gain in OS. The ERG had concerns that this assumption may not be reasonable and was not adequately supported by evidence. Overall, the ERG noted that the precise modelling approach adopted is not important provided that the selected model structure does not impose inappropriate restrictions on the synthesis of evidence and the model-predicted outcomes are credible. However, the company's model-predicted OS did not provide a good visual fit to the observed Kaplan-Meier curve for ibrutinib: rather, the model over-estimated the OS for ibrutinib up to around 15.6 months and subsequently under-predicted OS beyond this timepoint (Figure 1). This suggested that the modelled OS trajectory for the ibrutinib group had likely been underestimated. No equivalent data were available for the R-chemo group of the model, hence the accuracy of the OS prediction in this group is unknown. As a consequence of these issues, the ERG had concerns regarding the credibility of the company's results. The ERG noted that an alternative modelling approach, for example, the partitioned survival approach applied in a previous model of

ibrutinib for MCL [24], would have allowed for the inclusion of the available OS data for ibrutinib within the model and may have produced more robust predictions.

Figure 1: Comparison of observed and predicted overall survival for ibrutinib group
[INSERT FIGURE 1 HERE]

3.3.1.2 Issues surrounding use of parametric survival modelling

The ERG highlighted several concerns regarding the company's parametric survival modelling of data on pre-progression-mortality, PFS, and PPS. Within the CS [6], only a limited set of candidate survivor functions (exponential, Weibull, log normal and log logistic) were fitted to the available PFS data for ibrutinib. Other survivor functions, for example, the Gompertz, the generalised gamma, the gamma and the generalised F models should have been considered. Following clarification, the company fitted Gompertz and generalised gamma functions to the PFS data, but noted that both produced clinically implausible projections [10]. In addition, within the CS, both pre-progression mortality and PPS were assumed to be subject to a constant hazard rate. Following clarification, the company provided some evidence that the exponential model provided a reasonable fit to the pre-progression mortality data. The company also subsequently fitted alternative parametric functions to the PPS data [10], however, given that the model included a structural assumption that the PPS hazard is constant, these alternative functions could not be incorporated into the model, hence their impact on the ICER is unknown.

3.3.1.3 Methods for modelling time to treatment discontinuation (or death)

The company fitted exponential, Weibull, log logistic and log normal survivor functions to data on TTD/D from the pooled ibrutinib dataset [6]; the Weibull model was selected for use in the base case analysis. However, all fitted curves considerably overestimated the probability of being alive and on treatment beyond the final observed datapoint within the ibrutinib dataset. Based on the company's fitted Weibull model, the probability that a patient would still be receiving treatment at 50 months was approximately 7%, whilst the Kaplan-Meier curve indicated that all patients had discontinued by around 32 months. Consequently, the modelled drug costs for ibrutinib, and therefore the ICER, were overestimated. The ERG considered that whilst the Kaplan-Meier curve is most uncertain in its tail, the best estimate of the cumulative survival probability for TTD/D would be obtained from the observed Kaplan-Meier curve rather than a parametric model which does not provide a good fit to those data. The ERG also noted that if censoring was truly random, additional data collection over a longer follow-up should not produce a systematic shift in the Kaplan-Meier curve.

3.3.1.4 Uncertainty surrounding relative effectiveness of ibrutinib versus currently used treatments

Given the absence of any head-to-head trials of ibrutinib versus R-chemo in patients with R/R MCL, the company's economic analysis was hinged on an indirect comparison informed by the RAY study

[19], the OPTIMAL trial [14] and the HMRN audit [17]. ERG highlighted a number of limitations and uncertainties in the company's indirect comparison:

- The treatment effect for ibrutinib versus R-chemo, and the associated uncertainty around this estimate, could have been more meaningfully synthesised using a random effects NMA.
- The OPTIMAL trial [14] involved only single-agent chemotherapy, however, with the exception of cytarabine, all options included in the final NICE scope relate to combination chemotherapy regimens [9].
- The HR for PFS taken from the HMRN audit did not specifically relate to patients with R/R disease, did not differentiate between chemotherapy regimens, and reflected only those patients achieving response [17]. In addition, since this study is not an RCT, the HR may be subject to confounding due to differences in patient characteristics.
- The ERG's clinical advisors suggested that rituximab plus bendamustine (R-bendamustine), rather than R-CHOP, is likely to represent the main comparator for ibrutinib, although other R-chemo regimens may be considered and clinical outcomes for R-bendamustine and R-CHOP are likely to be similar.
- Pre-progression mortality for patients receiving R-chemo is not included as an outcome in the indirect comparison, but is instead assumed to be equal to the rate observed in the TEM arm of the RAY trial [19]. Given the absence of evidence, the validity of this assumption is unclear.
- The indirect comparison used in the health economic model is restricted to the outcome of PFS. The ERG noted that the HR for OS for R-chemo versus chemotherapy could have instead been derived using data from the RCT reported by Forstpointner *et al* [16].

In light of these issues, the ERG noted that any estimate of treatment effect for ibrutinib relative to R-chemo derived from the available evidence base will be subject to considerable uncertainty.

3.3.1.5 Additional concerns identified by the ERG

The ERG noted a number of other concerns regarding the company's economic analysis. These included: (i) the inappropriate exclusion of costs of subsequent therapies beyond progression; (ii) the use of blended comparison of R-chemo options with equal efficacy; (iii) an assumption of an HRQoL decrement associated with R-chemo for patients which is sustained beyond discontinuation, and (iv) uncertainty and risk of confounding in the company's subgroup analyses based on number of LOTs.

3.4 Additional work undertaken by the ERG

The ERG undertook two sets of exploratory analyses – "Set A" and "Set B". Exploratory analysis Set A was undertaken using the company's Markov model. The ERG's preferred analysis involved using the HR for PFS derived from ERG's random effects NMA, applying the observed Kaplan-Meier curve

for TTD/D for the ibrutinib group, and truncating the R-chemo disutility following treatment discontinuation. Based on this scenario, the probabilistic ICER for ibrutinib versus R-CHOP was £63,340 per QALY gained (Table 1). The ICERs for all other analyses based on this ERG-preferred model were greater than £59,952 per QALY gained. Within the one prior LOT subgroup, the ICER for ibrutinib versus R-CHOP was estimated to be £44,711 per QALY gained.

Exploratory analysis Set B involved converting the company's model structure to adopt a partitioned survival approach. This entailed using the available data for PFS and OS for ibrutinib from the pooled dataset and the estimation of treatment effects on PFS and OS derived from the ERG's NMAs. These analyses suggested that irrespective of whether the rituximab effect is estimated using data reported by Forstpointner *et al* [16], the HRMN audit [17], or both, ibrutinib is expected to be dominated. Importantly, the ERG noted that this economic conclusion was likely to be a consequence of problems in robustly estimating treatment effects for OS given the weaknesses in the available evidence.

3.5 Conclusion of the ERG report

As R/R MCL is a relatively rare disease, few real-world data are available. Only three studies of ibrutinib in R/R MCL patients were identified, and these did not reflect treatment pathways relevant to current clinical practice in England. Based on the ERG's additional analyses, ibrutinib is associated with a slower rate of disease progression compared with R-chemo (random effects HR=0.27, 95% CrI 0.06 to 1.26), although the result is inconclusive as it did not reach statistical significance at the 95% level. The estimated median HRs for OS for ibrutinib versus R-chemo ranged from 0.98 to 1.96, depending on the data source used for the rituximab arm of the network. This illustrates the high level of uncertainty for this comparison. Based on analyses of the company's model undertaken by the ERG (including the company's original PAS), the ICER for ibrutinib versus R-chemo is likely to be greater than £59,952 per QALY gained in the overall R/R MCL population. The cost-effectiveness profile of ibrutinib appears to be improved in the one prior LOT subgroup, but may be subject to confounding due to the *post hoc* definition of the subgroup and bias due to the poor fit of the Weibull function used to model PFS.

4. METHODOLOGICAL ISSUES

The ERG noted problems relating to the robustness of the indirect comparison for OS. The ERG considers that a balance exists in that the company's PFS-based Markov model makes a number of restrictive structural assumptions which lead to a poor model fit to the available OS data for ibrutinib, whilst the ERG's partitioned survival analysis (exploratory analysis Set B) provided a better fit to the OS data but involved using the outputs of a highly uncertain NMA.

5. NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE (NICE) GUIDANCE

The appraisal committee reviewed the data available on the clinical and cost effectiveness of ibrutinib, having considered evidence on the nature of R/R MCL and the value placed on the benefits of ibrutinib by people with the condition, those who represent them, and clinical experts. It also took into account the effective use of NHS resources.

The first ACD (published August 2016) did not recommend ibrutinib within its marketing authorisation for the treatment of adult patients with R/R MCL [25]. The AC noted that the ICERs presented by the company, which incorporated the original PAS for ibrutinib, were substantially above the range normally considered a cost-effective use of NHS resources. Following the first ACD, the company submitted an additional analysis of the HMRN dataset and updated the PAS for ibrutinib. The AC considered that this new analysis provided some reassurance regarding the company's modelling method, but noted that the ICER was higher than the committee could accept as a cost-effective use of NHS resources.

Following the second ACD, the company submitted a response which included a newer data-cut from the RAY study and requested routine commissioning for ibrutinib within the one prior LOT subgroup, based on an ICER of £49,849 per QALY gained. In December 2017, NICE published its FAD [3], which made the following recommendation: *"Ibrutinib is recommended as an option for treating relapsed or refractory mantle cell lymphoma in adults, only if:*

- *they have had only 1 previous line of therapy and*
- *the company provides ibrutinib with the discount agreed in the commercial access agreement with NHS England."*

5.1 Consideration of clinical effectiveness and cost-effectiveness issues

This section discusses the key issues considered by the AC. The full list can be found in the FAD [3].

5.1.1 Clinical trial evidence

The AC noted that the evidence on the clinical effectiveness of ibrutinib came from one RCT (RAY) and two single-arm studies (SPARK and PCYC-1104). It considered that RAY is not strictly relevant to NHS practice because TEM is not routinely used in the UK and noted the absence of trials comparing ibrutinib with any comparator defined in the NICE scope. The committee concluded that the studies were of a reasonable quality but were limited by the lack of a comparison against a treatment used in UK clinical practice. The committee noted that at median follow-up of 20 months in RAY, median PFS was statistically significantly longer for ibrutinib compared with TEM (14.6 months compared with 6.2 months; HR=0.43; 95% CI 0.32 to 0.58; p<0.0001). At the time of the first AC meeting, the OS data

from RAY were immature and median OS had not yet been reached in the ibrutinib arm. The committee understood that the OS results could be confounded as 23% of patients in the TEM arm switched to subsequently receive ibrutinib and the use of subsequent anticancer systemic therapies differed between treatment groups. Following consultation on the second ACD, the committee considered updated data from RAY provided by the company; they noted that that the updated results were consistent with the earlier data and that median OS had now been reached (30.3 months for ibrutinib compared with 23.5 months for TEM; HR=0.74; 95% CI 0.54 to 1.02). The committee concluded that the results from RAY suggest that ibrutinib significantly improves PFS compared with TEM, but that the OS benefits remain uncertain.

5.1.2 Indirect comparison

The committee understood the company's indirect comparison using data from RAY, OPTIMAL and the HMRN audit resulted in an HR for PFS for ibrutinib versus R-chemo of 0.28. The committee acknowledged the limitations of the indirect comparison and noted that the ERG did not agree with the company's two-stage approach to estimating treatment effects. The committee noted that because of concerns about the evidence used to inform the indirect comparisons, the ERG considered that the results of both the company's analyses and the ERG's analyses should be interpreted with caution. The committee concluded that there is considerable uncertainty associated with the indirect comparisons and that the benefit of ibrutinib compared with R-chemo is unclear, although it accepted that the available evidence and experience from clinical practice strongly suggest that ibrutinib is more effective.

5.1.3 Subgroups

The committee discussed the efficacy results for subgroups of patients, based on the number of previous LOTs. It noted that the results suggest greater efficacy in patients who had ibrutinib after only one previous LOT, compared with two or more previous LOTs. The clinical expert also stated that ibrutinib is particularly beneficial after the first relapse. The committee considered the updated RAY data and noted that these provide further evidence of a greater benefit of ibrutinib when taken after only one previous LOT. The committee understood that the data were potentially confounded by patients in the TEM arm switching to ibrutinib and that the subgroups were defined *post hoc*. However, the committee noted responses to the ACD from professional groups that stated that evidence from clinical practice supports the RAY results, and that earlier use of ibrutinib in R/R disease is the most beneficial. The committee concluded that the evidence from RAY and clinical experience suggests that ibrutinib is most effective in people who have had only one previous LOT.

5.1.4 Company's economic analysis

The committee noted that the company had developed a Markov model comparing ibrutinib with R-chemo, comprising three states (pre-progression, post-progression and death), and that this approach had been used in previous appraisals. The committee was aware that OS data from the ibrutinib studies were not directly extrapolated but were modelled using PFS and PPS data from the pooled ibrutinib dataset. The committee considered that the company's approach is appropriate given the immaturity of the OS data available.

The committee considered the ERG's critique of the company's model. It noted the ERG's comments that the company's Markov approach imposed structural constraints, which did not make the best use of the trial data on survival, and that the OS predicted by the model did not provide a good visual fit to the observed Kaplan-Meier OS curve from the trials. The committee understood that the ERG had explored the effect of using a partitioned survival approach in their exploratory analysis (Set B), but was concerned that this resulted in efficacy estimates for R-chemo that were higher than those for ibrutinib, giving higher QALY gains for R-chemo than ibrutinib which were implausible. The committee concluded that the results of the partitioned survival analysis were not clinically plausible, acknowledging the ERG's comments that they are associated with major uncertainty because they rely on the outputs of a highly uncertain meta-analysis.

The committee re-examined the company's Markov approach, which it considered led to more plausible results. The committee noted that in the company's base-case analysis which incorporated the updated PAS, the ICER for ibrutinib versus R-chemo was £62,650 per QALY gained. The committee also noted that all but one of the company's scenario analyses produced ICERs which were above £59,000 per QALY gained. In this one scenario, the company applied an HR to post-progression survival for R-chemo, which was adjusted to be as close as possible to the anticipated survival based on the results of the HMRN audit. This resulted in an ICER of £49,849 per QALY gained. However, the committee understood that time-to-event estimates for PFS and PPS for ibrutinib were taken from the one prior LOT subgroup, and therefore that the analysis reflects this subgroup.

The committee noted that the ERG's exploratory analyses using the company's model (Set A) made adjustments to some of the parameter values in the company's model which mostly resulted in lower ICERs than those presented by the company. However, the committee was minded not to accept the results of the ERG's preferred base case for Set A because these represented the extreme (lowest) end of the ERG's wide estimate of possible ICERs, depending on the model and parameters used. The committee concluded that the ICERs presented by the company for the whole population of people with R/R MCL, incorporating the updated PAS for ibrutinib, were above the range normally considered a cost-effective use of NHS resources (£20,000 to £30,000 per QALY gained). The committee recognised

that ibrutinib has several benefits which may not be reflected in the QALY calculations (oral administration, manageable adverse reactions and low toxicity); however, it did not consider that these would be sufficient to lower the ICER for the whole population to within the range normally considered to be cost-effective.

Noting its conclusion that trial evidence and clinical experience suggest that ibrutinib is most effective in people who have had only one prior LOT, the committee considered whether ibrutinib could be considered cost-effective in this subgroup of patients. It noted that the company's ICER of £49,849 per QALY gained may be a conservative estimate because updated trial data from RAY suggest that the model underestimates survival for this subgroup. The committee also noted that overall survival in RAY may have been confounded by the switching of patients in the TEM arm to the ibrutinib arm. The committee concluded that the most plausible ICER in this group of patients is likely to be lower than the company's estimate of £49,848 per QALY gained.

The committee noted that the OS estimates presented for people with R/R MCL ranged from 5.2 months to 9.7 months. It also accepted that there is enough evidence to indicate that ibrutinib offers an extension to life of at least an additional 3 months, compared with current NHS treatment. On this basis, the committee concluded that ibrutinib met all the criteria to be considered a life-extending end-of-life treatment.

6. APPRAISAL COMMITTEE'S KEY CONCLUSION

The committee concluded that the most plausible ICER for the one prior LOT subgroup is likely to be lower than the company's estimate of £49,848 per QALY gained.

Author contributions

Emma Simpson and Eva Kaltenthaler summarised and critiqued the clinical effectiveness data reported within the company's submission. Mark Clowes critiqued the company's search strategy. Jean Hamilton critiqued the statistical analyses undertaken by the company and undertook additional analyses. Paul Tappenden and Daniel Pollard critiqued the health economic analysis submitted by the company and undertook the ERG's exploratory analyses. David Meiklejohn and Nick Morley provided clinical advice to the ERG throughout the project. All authors were involved in drafting and commenting on the final report. Paul Tappenden acts as the guarantor of the manuscript. This summary has not been externally reviewed by PharmacoEconomics.

Compliance with ethical standards

Funding

This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment Programme (project no. 14/177/09). See the HTA programme website for further project information (http://www.hta.ac.uk). This summary of the ERG report was compiled after NICE issued the FAD. All authors have commented on the submitted manuscript and have given their approval for the final version to be published. The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

Conflicts of interest

Paul Tappenden, Emma Simpson, Jean Hamilton, Daniel Pollard, Mark Clowes, Eva Kaltenthaler, David Meiklejohn and Nick Morley declare no financial conflicts of interest.

REFERENCES

1. National Institute for Health and Care Excellence. Guide to the processes of technology appraisal. NICE: London; 2014. Available from: https://www.nice.org.uk/process/pmg9/chapter/foreword [accessed 01/04/2018].
2. Tappenden P, Simpson E, Sanderson J, Pollard D, Clowes M, Kaltenthaler E, *et al*. Ibrutinib for treating relapsed or refractory mantle cell lymphoma: A Single Technology Appraisal. University of Sheffield: Sheffield; 2016. Available from: https://www.nice.org.uk/guidance/ta502/documents/committee-papers [accessed 04/03/2018].
3. National Institute for Health and Care Excellence. Ibrutinib for treating relapsed or refractory mantle cell lymphoma: Final Appraisal Determination. NICE: London; 2017. Available from https://www.nice.org.uk/guidance/ta502/documents/final-appraisal-determination-document [accessed 01/05/2018].
4. National Institute for Health and Care Excellence. Ibrutinib for treating relapsed or refractory mantle cell lymphoma: Technology Appraisal Guidance. NICE: London; 2017. Available from: https://www.nice.org.uk/guidance/ta502 [accessed 01/05/2018].
5. National Institute for Health and Care Excellence. Ibrutinib for treating relapsed or refractory mantle cell lymphoma. NICE: London; 2017. https://www.nice.org.uk/guidance/ta502 [accessed 01/06/2018].
6. Janssen Ltd. Ibrutinib for the treatment of relapsed or refractory mantle cell lymphoma. Company's evidence submission to the National Institute for Health and Care Excellence. Janssen: High Wycombe; 2016. Available from: https://www.nice.org.uk/guidance/ta502/documents/committee-papers [accessed 04/03/2018].
7. McKay P, Leach M, Jackson R, Cook G, Rule S. Guidelines for the investigation and management of mantle cell lymphoma. *British Journal of Haematology* 2012;159:405-426.
8. Dreyling M, Thieblemont C, Gallamini A, Arcaini L, Campo E, Hermine O, *et al*. ESMO consensus conferences: guidelines on malignant lymphoma. Part 2: marginal zone lymphoma, mantle cell lymphoma, peripheral T-cell lymphoma. *Annals of Oncology* 2013;24(4):857-877.
9. National Institute for Health and Care Excellence. Ibrutinib for treating relapsed or refractory mantle cell lymphoma: Final scope. NICE: London; 2016. Available from: https://www.nice.org.uk/guidance/ta502/documents/final-scope [accessed 01/04/2018].
10. Janssen Ltd. Ibrutinib for treating mantle cell lymphoma. Company's response to clarification questions from the ERG. Janssen: High Wycombe; 2016. Available from: https://www.nice.org.uk/guidance/ta502/documents/committee-papers [accessed 04/03/2018].
11. Dreyling M, Jurczak W, Jerkeman M, Silva RS, Rusconi C, Trneny M, *et al*. Ibrutinib versus temsirolimus in patients with relapsed or refractory mantle-cell lymphoma: an international, randomised, open-label, phase 3 study. *Lancet* 2016;387(10020):770-8.
12. Wang ML, Rule S, Martin P, Goy A, Auer R, Kahl BS, *et al*. Targeting BTK with ibrutinib in relapsed or refractory mantle-cell lymphoma. *New England Journal of Medicine* 2013;369(6):507-516.
13. Wang M, Goy A, Martin P, Ramchandran R, Alexeeva J, Popat R. Efficacy and safety of single-agent ibrutinib in patients with mantle cell lymphoma who progressed after bortezomib therapy. *56th American Society of Hematology (ASH) Annual Meeting and Exposition.* San Francisco, US; 6-9 December 2014.
14. Hess G, Herbrecht R, Romaguera J, Verhoef G, Crump M, Gisselbrecht C, *et al*. Phase III study to evaluate temsirolimus compared with investigator's choice therapy for the treatment of relapsed or refractory mantle cell lymphoma. *Journal of Clinical Oncology* 2009;27(23):3822-3829.

15. Carter GC, Liepa A, Zimmerman AH, Morschhauser F. Validation of the Functional Assessment of Therapy—Lymphoma (FACT-LYM) in patients with relapsed/refractory mantle cell lymphoma. (Abstract 2376). *Blood* 2008;112:828.

16. Forstpointner R, Dreyling M, Repp R, Hermann S, Hänel A, Metzner B, *et al*. The addition of rituximab to a combination of fludarabine, cyclophosphamide, mitoxantrone (FCM) significantly increases the response rate and prolongs survival as compared with FCM alone in patients with relapsed and refractory follicular and mantle cell lymphomas: results of a prospective randomized study of the German Low-Grade Lymphoma Study Group. *Blood* 2004;104(10):3054-3071.

17. Haematological Malignancy Research Network. Clinical management and outcome in mantle cell lymphoma - Version 1.4. HMRN: York; 2016.

18. Janssen Research and Development. Data on File – Imbruvica (ibrutinib) – Pooled analysis [data held on file]. 2016. Janssen: High Wycombe; 2016.

19. Janssen Research and Development. RAY (MCL3001) Clinical Study Report. Janssen: High Wycombe; 2015.

20. Janssen Research and Development. SPARK (MCL2001) Clinical Study Report. Janssen: High Wycombe; 2014.

21. Haymarket Media Group. Monthly Index of Medical Specialties (MIMS) 2016. Available from: https://www.mims.co.uk/.

22. Commercial Medicines Unit. Drugs and pharmaceutical electronic market information (eMit). Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit2015.

23. Department of Health. Reference Costs 2014/15. DH: London; 2015. Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2014-to-2015 [accessed 01/05/2016].

24. Peng S, Sorensen S, Pan F, Dorman E, Sun S, Van Sanden S, *et al*. Simulation model of ibrutinib in treatment of relapsed or refractory mantle cell lymphoma (MCL*). Value in Health* 2014;17(7):A620.

25. National Institute for Health and Care Excellence. Ibrutinib for treating relapsed or refractory mantle cell lymphoma: Appraisal Consultation Document. NICE: London; 2016. Available from: https://www.nice.org.uk/guidance/ta502/documents/appraisal-consultation-document [accessed 01/05/2018].

26. Lachaine J, Beauchemin C, Mathurin K, Aissa F. Cost-effectiveness of bendamustine+rituximab versus fludarabine+rituximab in the treatment of relapsed indolent non-Hodgkin's and mantle cell lymphomas in Canada. *Value in Health* 2013;12(A141).

27. Yoong K, Attard C, Jivraj F, Sehn L. Cost-effectiveness analysis of bortezomib in relapsed mantle cell lymphoma patients in Canada. *Value in Health* 2009;12(A273).

Figure 1: Comparison of observed and predicted overall survival for ibrutinib group

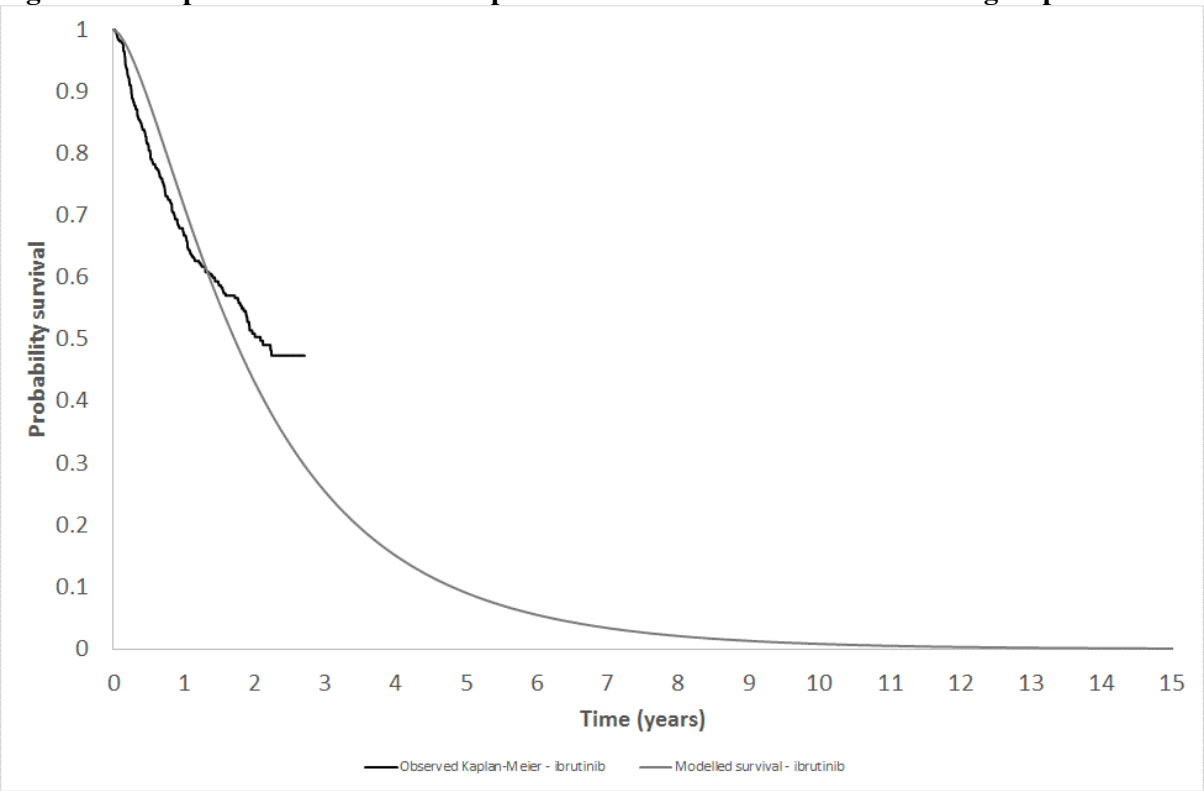

Table 1: Summary of key results from the CS and the ERG report (using the original PAS for ibrutinib in the MCL indication)

Scenario	ICER (per QALY gained) Ibrutinib versus R-chemo
Company's economic analysis	
Company's base case (probabilistic)	£76,014
Company's most favourable DSA (Assuming R-chemo post-progression mortality probability = 0.27 (to reflect HMRN median OS of 8.4 months) using the 1 prior LOT subgroup)	£59,345
Company's least favourable DSA (PFS modelled using exponential distribution)	£80,296
Subgroup analysis – 1 prior LOT subgroup	£65,977
ERG exploratory analyses, Set A – based on the company's Markov model	
ERG exploratory analysis A1: HR for PFS derived from ERG's random effects NMA	£75,094
ERG exploratory analysis A2: TTD/D for ibrutinib group based on Kaplan-Meier curve	£61,472
ERG exploratory analysis A3: Truncation of R-chemo disutility following treatment discontinuation	£77,111
Exploratory analysis A4: ERG's preferred analysis using the company's model (combining ERG analyses A1, A2 and A3, probabilistic)	£63,340
Exploratory analysis A5i: Utilities for progression-free and post-progression based on Lachaine *et al*[26], based on the ERG's preferred analysis	£60,417
Exploratory analysis A5ii: Utilities for progression-free and post-progression based on Yoong *et al*[27], based on the ERG's preferred analysis	£59,952
Exploratory analysis A6i: Cost-effectiveness of ibrutinib versus chemotherapy for rituximab-resistant patients (rituximab cost set equal to zero), based on the ERG's preferred analysis	£69,054
Exploratory analysis A6ii: Cost-effectiveness of ibrutinib versus chemotherapy for rituximab-resistant patients (Cost of rituximab set to zero and PFS HR=0.19), based on the ERG's preferred analysis	£64,727
Exploratory analysis A7: Ibrutinib versus R-CHOP in the 1 prior LOT subgroup using the ERG's preferred analysis, based on the ERG's preferred analysis	£44,711
ERG exploratory analyses, Set B – based on the ERG's partitioned survival model and NMAs for PFS and OS	
Rituximab effect informed by Forstpointner *et al*[16]	Dominated
NMA rituximab effect informed by HMRN[17]	Dominated
NMA rituximab effect informed by Forstpointner *et al*[17]	Dominated

ICER – incremental cost-effectiveness ratio; QALY – quality-adjusted life year; DSA – deterministic sensitivity analysis; HMRN – Haematological Malignancy Research Network; LOT – line of therapy; HR – hazard ratio; OS – overall survival; PFS – progression-free survival; TTD/D – time to treatment discontinuation or death; R-CHOP - rituximab, cyclophosphamide, doxorubicin and vincristine; NMA – network meta-analysis; ERG – Evidence Review Group