Cost Effectiveness of Rituximab for Non-Hodgkin’s Lymphoma
A Systematic Review

Philipp W.P. Auweiler,1 Dirk Müller,1 Stephanie Stock1 and Andreas Gerber2

1 Institute for Health Economics and Clinical Epidemiology, The University Hospital of Cologne (AöR), Cologne, Germany
2 Institute for Quality and Efficiency in Health Care, Cologne, Germany

Abstract

Background: The monoclonal antibody rituximab has shown clinical effectiveness in combination with chemotherapy for the treatment of non-Hodgkin’s lymphoma (NHL) in several randomized controlled studies. Rituximab maintenance therapy is associated with significant improvement in progression-free and overall survival in patients with NHL. However, treatment with rituximab causes considerable costs for healthcare systems.

Objective: This article provides an overview of economic evaluations of rituximab and appraises their methodological quality.

Methods: A systematic literature search of cost-effectiveness studies on rituximab was carried out in nine electronic databases: MEDLINE, EMBASE, Cochrane Database of Systematic Reviews (CDSR), the German Agency of Health Technology Assessment (DAHTA) database, German Institute for Quality Improvement (DIQ)-Literatur, DIQ-Projekte, Database of Abstracts of Reviews of Effects (DARE), Health Technology Assessments (HTA) database and Sozialmedizin (SOMED) [languages: English, German, Dutch, French, Spanish and Italian; publication period: 1998 to 2010]. Based on pre-specified inclusion criteria, cost-effectiveness studies were identified that compared standard chemotherapy with standard chemotherapy plus rituximab in patients with a subtype of NHL. The methodological quality of the studies was assessed using a quality checklist.

Results: Fourteen economic evaluations from seven different countries were included in the review. All economic evaluations reported incremental cost-effectiveness ratios (ICERs) for the add-on therapy with rituximab that were below the country-specific thresholds. The studies differed significantly in their characteristics and methodological rigour. Most studies lacked transparency regarding identification and justification of data. In several studies, the rationale for the model structure was not described appropriately.

Conclusion: Adding rituximab to standard chemotherapy is considered a cost-effective treatment option for NHL. However, the results of the analyses should be interpreted with caution due to methodological limitations.
1. Introduction

Non-Hodgkin’s lymphoma (NHL) compromises a group of malignancies of the lymphoid system. Its incidence has doubled over the last two decades in the US as well as in most Western countries.[1] It is estimated that 65,540 people (35,380 men and 30,160 women) were diagnosed with NHL and 20,210 men and women died of NHL in 2010 in the US.[2] Among all estimated US cancer deaths in 2010, NHL was the reason for 4% of cancer deaths among men and women.[3]

Treatments with single agent or combination chemotherapy such as cyclophosphamide, doxorubicin, vincristine and prednisone (CHOP) have achieved significant results with regard to complete or partial remission.[4,5] In the late 1990s, rituximab emerged as a new treatment. Rituximab is a monoclonal antibody targeted against CD20-positive lymphoma cells, involving cell-mediated cytotoxicity and apoptosis.[6] The clinical effectiveness of rituximab in combination with chemotherapy in several subtypes of NHL, including follicular lymphoma (FL), diffuse large B-cell lymphoma (DLBCL) and low grade B-cell lymphoma, has been demonstrated by several randomized controlled trials (RCTs).[6-12] Rituximab maintenance therapy has also achieved significant improvements in progression-free survival (PFS) and overall survival in patients with NHL.[6,13-15] Rituximab obtained market authorization in Europe for patients with relapsed/refractory FL.[16] Although newer drugs are available for the treatment of NHL (e.g. ibrutinumab-tiuxetan), rituximab is recommended in the treatment of NHL patients in many countries.

Given the fact that the application of rituximab causes extensive costs in healthcare, several cost-effectiveness studies and/or cost-utility studies on rituximab have been carried out.[16-37] For decision makers it is essential to assess the reliability of these studies because economic evaluations, by nature, tend to differ in the data used for the analyses, the model structure, or in their overall methodological quality.

To provide an overview of the cost effectiveness of rituximab and the quality of the evidence, the aim of the present systematic literature review was to systematically appraise and review all published economic evaluations between 1998 and 2010 on rituximab for the treatment of subtypes of NHL in patients aged >18 years. Upon the results of the assessment of the methodological rigour of the studies, conclusions were drawn with regard to the trustworthiness of the results and their recommendation as basis for healthcare policy decisions.

2. Methods

2.1 Literature Search

A systematic literature search of cost-effectiveness studies on rituximab was carried out by one reviewer in nine electronic databases: MEDLINE, EMBASE, Cochrane Database of Systematic Reviews (CDSR), the German Agency of Health Technology Assessment (DAHTA) database, German Institute for Quality Improvement (DIQ)-Literatur, DIQ-Projekte, Database of Abstracts of Reviews of Effects (DARE), Health Technology Assessment (HTA) database and Sozialmedizin (SOMED). Publications in English, German, Dutch, French, Spanish and Italian from the year 1998 to 2010 (August) were selected. Using a comprehensive number of search terms, the
search strategy combined different economic terms with the terms for NHL and the intervention to be evaluated. For instance, the search in MEDLINE was carried out combining the Medical Subject Headings (MeSH) ‘Lymphoma, Non-Hodgkin’ AND ‘Cost and Cost Analysis’ AND ‘Rituximab’ (for more details, see the technical appendix in the Supplemental Digital Content [SDC], http://links.adisonline.com/PCZ/A143). Furthermore, a manual reference search was conducted in order to identify additional studies.

2.2 Inclusion Criteria

The following inclusion criteria were applied: cost-effectiveness or cost-utility studies with estimates of incremental cost-effectiveness ratios (ICERs); data on clinical effectiveness based on RCTs that compared a standard chemotherapy regime with a standard chemotherapy regime plus rituximab; patients with a subtype of NHL (older than 18 years); and availability of the full text of the identified article.

2.3 Assessment of Methodological Quality

In order to assess the methodological quality of the cost-effectiveness studies, the checklist of Philips et al.[38] was applied. This checklist is a tool to detect the potential strengths and shortcomings of pharmacoeconomic modelling studies. The checklist addressed the following dimensions of quality: statement of decision problem/objective (S1), statement of scope/perspective (S2), rationale for structure (S3), structural assumptions (S4), strategies/comparators (S5), model type (S6), time horizon (S7), disease states/pathways (S8), cycle length (S9), data identification (D1), pre-model data analysis (D2), baseline data (D2a), treatment effects (D2b), quality-of-life (QOL) weights (D2d), data incorporation (D3), assessment of uncertainty (D4), internal consistency (C1) and external consistency (C2).

To assess the methodological quality of the studies included in this review, we used a rating system. For each criterion of the checklist that was achieved, a positive point (+) was allocated. Results are reported as scores. The checklist, as well as the detailed description of the allocation of the scores, is provided in the technical appendix in the SDC.

The economic studies were rated by two trained reviewers (DM, PA). Any discrepancies between the reviewers were clarified on a face-to-face basis. In addition, the RCTs (pivotal trials) from which data on clinical effectiveness were derived were checked using the Jadad quality checklist (see the SDC). A Jadad score of two points or higher was regarded as methodologically sound.[17] Concerning the outcomes of cancer progression trials, there are serious practical objections to clinical blinding and ethical objections to patient blinding.[17]

3. Results

After screening of titles and abstracts, 27 cost-effectiveness studies were obtained. Based on the full texts of these articles, 13 were excluded from the review due to different reasons (figure 1). The remaining 14 analyses were conducted in seven different countries (the US = 3; the UK = 3; Italy = 3; France = 2; the Netherlands = 1; Spain = 1; Sweden = 1).

![Flow chart of the search and assessment of the literature.](image)
3.1 Methodological Assessment

The quality score ranged from 6 to 49 out of a possible 57 points for the economic evaluations. While one study reached a score of 49 points,\[^{30}\] five studies reached a score of 30 or more points.\[^{16,17,22,35,36}\] The detailed results of the methodological assessment are summarized in table I.

Assessing dimension S2 (statement of scope/perspective), five studies\[^{17,21,22,32,35}\] failed to include model inputs according to the chosen perspective of the analysis and/or did not state and justify the scope of the model. Poor performance was found in dimension S3 (rationale for structure) due to a lack of reporting the chosen model structure\[^{16,20,21,32,34,36}\] or competing theories regarding the model structure,\[^{16-18,20-22,29,31-36}\] and lack of justification of the causal relationships described by the model.\[^{18,20,21,29,31-36}\]

Except for one analysis,\[^{30}\] it was unclear if all feasible and practical options concerning the stated decision problem had been evaluated (dimension S5: strategies/comparators). Concerning dimension S9 (cycle length), only in four of 12 health transition models\[^{16,30,35,36}\] was the cycle length defined and justified in terms of the natural history of disease.

The items concerning the dimension D1 (data identification) were not achieved by quite a few evaluations.\[^{18,20,21,29,31-35}\] Common shortcomings were lack of systematic methods to identify the most appropriate data, lack of quality assessments of selected data, and lack of describing the sources of expert opinions and the methods of reaching agreement. Half-cycle corrections (dimension D2a: baseline data) were only applied to three of 12 health transition models.\[^{30,35,36}\] In addition, poor performance was detected in dimension D2d (QOL weights) because the methods of the derivation for utility values were not justified or were unclear in many of the evaluations.

For dimension D4, none of the evaluations provided estimations of uncertainty of the analyses that included all four principal types of uncertainty (i.e. methodological, structural and parameter uncertainty, and heterogeneity). Three studies only reported results from a deterministic sensitivity analysis.\[^{32-34}\]

3.2 Results of Economic Evaluations

All pharmacoeconomic evaluations reported ICERs that were below the country-specific thresholds. With respect to evaluations that presented base-case results in €, the cost per life-year gained (LYG) ranged from €7612 to €16 493. Results from cost-utility analyses ranged from €8729 to €29 976 per QALY. Estimations in SUS ranged from SUS12 304/LYG to SUS17 504/LYG, and from SUS19 297/QALY to SUS28 565/QALY, respectively. Base-case results in £ ranged from £8532/LYG to £9774/LYG, and from £6503/QALY to £26 000/QALY, respectively. In all analyses, rituximab was considered as cost effective compared with the reference scenario. With regard to the degree of uncertainty of the results, some studies were affected by uncertainty around cost-effectiveness point estimates.

3.3 Study Characteristics

In table II, a brief summary of the characteristics of the included studies is displayed. The characteristics were objective of the published study, setting/perspective, type of economic analysis, discount rates and ICERs. Additional information is presented in the following sections.

3.4 Clinical Effectiveness (Pivotal Trial)

Five studies\[^{17,18,20-22}\] used a phase III trial conducted by the Group d’Etude des Lymphomes de l’Adulte (GELA-LNH 98.5) with 399 patients older than 60 years with DLBCL as their source for clinical effectiveness. This trial scored two out of a possible five points in the Jadad score (see the Appendix in the SDC).

Hornberger et al.\[^{32}\] utilized clinical effectiveness data from the clinical trial conducted by Marcus et al.\[^{9}\] (Jadad score 2/5). Berto et al.,\[^{29}\] amongst others,\[^{16,30,31,34,35}\] relied on data on clinical effectiveness from the European Organization for Research and Treatment of Cancer (EORTC) 20981 trial.\[^{6}\] In this trial, 465 patients were randomly assigned to CHOP or rituximab plus CHOP [R-CHOP] (Jadad score 2/5). Ferrara and Ravasio\[^{33}\] made use of the clinical trial conducted by Pfreundschuh et al.\[^{11}\] This trial com-
## Table I. Methodological assessment

<table>
<thead>
<tr>
<th>Dimension of quality</th>
<th>Study</th>
<th>[31]</th>
<th>[33]</th>
<th>[32]</th>
<th>[22]</th>
<th>[21]</th>
<th>[20]</th>
<th>[17]</th>
<th>[34]</th>
<th>[18]</th>
<th>[19]</th>
<th>[16]</th>
<th>[26]</th>
<th>[49]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Structure (S)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1 Statement of decision problem/objective</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>S4 Structural assumptions</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>S7 Time horizon</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>S8 Disease states/pathways</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td><strong>Data (D)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D2 Premodel data analysis</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>D2a Baseline data</td>
<td>++</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>D3 Data incorporation</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>D4 Assessment of uncertainty</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>D4a Methodological</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>D4b Structural</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>D4c Heterogeneity</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td><strong>Consistency (C)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1 Internal consistency</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td><strong>Score</strong></td>
<td>22/57</td>
<td>21/57</td>
<td>12/57</td>
<td>31/57</td>
<td>20/57</td>
<td>24/57</td>
<td>34/57</td>
<td>65/7</td>
<td>26/57</td>
<td>25/57</td>
<td>31/57</td>
<td>35/57</td>
<td>39/57</td>
<td>49/57</td>
</tr>
</tbody>
</table>

a  See the ‘Methodological rigour’ section of the SDC, http://links.adisonline.com/PCZ/A143, for study scoring methodology. The score is the sum of all positive points that were allocated.

b  See the methodological checklist in the SDC for dimension of quality descriptions. Many of the dimensions of quality contained more than one attribute question.

NA = not applicable; SDC = Supplemental Digital Content; ? indicates that information provided in the article was unclear; + indicates that a positive point was given; − indicates that a point was not given.
<table>
<thead>
<tr>
<th>Study, year of value</th>
<th>Objective</th>
<th>Setting/perspective</th>
<th>Analysis type</th>
<th>Discounting</th>
<th>ICERs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groot et al.,[22] 2002</td>
<td>R-CHOP vs CHOP alone in DLBCL pts</td>
<td>Societal, Netherlands</td>
<td>CEA/CUA</td>
<td>Costs: 4% Effects: 4%</td>
<td>CEA: £14865/LYG for pts &lt;60 y; £16483/LYG for pts &gt;60 y CUA: £13983/QALY for pts &lt;60 y; £17933/QALY for pts &gt;60 y</td>
</tr>
<tr>
<td>Ferrara and Ravasio,[33] 2007</td>
<td>R-CHOP vs CHOP alone in DLBCL pts</td>
<td>Italian NHS</td>
<td>CEA</td>
<td>Costs: 3% Effects: 3%</td>
<td>(£) Dominant</td>
</tr>
<tr>
<td>Kasteng et al.,[31] 2007</td>
<td>R maintenance therapy vs observation after second-line therapy in pts with FL</td>
<td>Healthcare provider, Sweden</td>
<td>CEA/CUA</td>
<td>Costs: 3% Effects: 3%</td>
<td>CEA: £11200/LYG CUA: £12600/QALY</td>
</tr>
<tr>
<td>Best et al.,[20] 2003</td>
<td>R-CHOP vs CHOP alone in pts with DLBCL</td>
<td>French payer (French social security system)</td>
<td>CUA</td>
<td>Costs: 3% Effects: 3%</td>
<td>£29976/QALY</td>
</tr>
<tr>
<td>Knight et al.,[17] 2000</td>
<td>R-CHOP vs CHOP alone as first-line therapy for DLBCL</td>
<td>Societal, UK</td>
<td>CEA/CUA</td>
<td>Costs: 6% Effects: 1.5%</td>
<td>CEA: £8532/LYG for pts &lt;60 y; £9774/LYG for pts &gt;60 y CUA: £7533/QALY for pts &lt;60 y; £10596/QALY for pts &gt;60 y</td>
</tr>
<tr>
<td>Hayslip and Simpson,[34] 2006</td>
<td>Extended adjuvant R therapy vs observation for FL in second remission</td>
<td>US healthcare system</td>
<td>CUA</td>
<td>Costs: 3% Effects: 3%</td>
<td>$US19522/QALY</td>
</tr>
<tr>
<td>Berto et al.,[18] not reported</td>
<td>R-CHOP vs CHOP alone in pts with aggressive FL</td>
<td>Italian healthcare system</td>
<td>CEA/CUA</td>
<td>Costs: 6% Effects: 1.5%</td>
<td>CEA: £13717/LYG for pts &lt;60 y; £13372/LYG for pts &gt;60 y CUA: £12879/QALY for pts &lt;60 y; £13362/QALY for pts &gt;60 y</td>
</tr>
<tr>
<td>Berto et al.,[29] 2006</td>
<td>R maintenance therapy vs observation in relapsed/refractory FL pts following response to induction therapy with or without R</td>
<td>Italian public payer</td>
<td>CEA/CUA</td>
<td>Costs: 3.5% Effects: 3.5%</td>
<td>CEA: £9903/LYG CUA: £11097/QALY</td>
</tr>
<tr>
<td>Grupo de Farmacoeconomica del Linfoma Follicular,[35] 2006</td>
<td>R maintenance treatment vs observation in pts with FL</td>
<td>Spanish NHS</td>
<td>CEA/CUA</td>
<td>Costs: 3.5% Effects: 3.5%</td>
<td>CEA: £8493/LYG CUA: £9358/QALY</td>
</tr>
<tr>
<td>Deconinck et al.,[16] 2006</td>
<td>R maintenance therapy in FL</td>
<td>French payer (French social security system)</td>
<td>CEA/CUA</td>
<td>Costs: 3.5% Effects: 0%</td>
<td>CEA: €7612/LYG CUA: €8729/QALY</td>
</tr>
</tbody>
</table>

Continued next page
Comparing R-CHOP versus CHOP alone enrolled 824 patients age 18 through 60 years (Jadad score 2/5).

In the analysis of Ray et al.,[36] clinical effectiveness was based on four RCTs.[10,39-41] The evaluation that compared rituximab plus cyclophosphamide, vincristine and prednisone (R-CVP) with cyclophosphamide, vincristine and prednisone (CVP) alone was based on the trial conducted by Marcus et al. enrolling 321 patients (Jadad score 2/5). The trial by Herold et al.[40] compared rituximab plus mitoxantrone, chlorambucil and prednisolone (MCP) versus MCP alone and enrolled 358 patients (Jadad score 2/5). R-CHOP and CHOP alone were compared in the trial that was conducted by Hiddemann et al.[10] In this study, 428 patients were randomly assigned to the treatment arms (Jadad score 3/5). The evaluation comparing rituximab with cyclophosphamide, etoposide, doxorubicin, prednisolone plus interferon (CHVP-interferon) obtained clinical effectiveness from the trial conducted by Foussard et al.[41] that included 358 patients (Jadad score 2/5).

3.5 Model and Time Horizon

Five studies implemented a Markov model with three health states: progression free (PF), progressive disease (PD) and death (D).[16,29,31,35,36] The time horizon of these studies ranged from 10 years[35] to a patient’s lifetime.[36]

Two economic evaluations applied a Markov health transition model with three health states that was developed by the School for Health and Related Research (ScHARR) of the University of Sheffield.[17,22] The three health states were complete responder (CR) to treatment, non-responder and relapse from CRs (NR), and death. Both models calculated the ICER over a time horizon of 15 years.

Hornberger and Best[21] developed a Markov model with five health states over a time horizon of 5 years: event free, salvage, transplantation, end-of-life care and death. The Markov model of Hayslip and Simpson[34] comprised six health states: disease-free survivor, undergoing salvage treatment, subsequent remissions, refractory disease, transplantation and death. The time horizon was fixed at 5 years.
The economic analysis by Berto et al.\[18\] is based on a Markov model that evaluates costs and consequences over a period of 15 years. It defines five health states: start therapy, complete response, no response, progression and death. The Markov model by Hornberger et al.\[32\] includes three health states: time until progression or death, referred to as PFS; time after progression; and death with a time horizon of 30 years.

Ferrara and Ravasio\[33\] presented a pharmacoeconomic evaluation that did not implement a health transition model but a decision-analytic model with a tree structure. The model provided estimations over a time horizon of 3 years.

Best et al.\[20\] used a model with a time horizon of 15 years, which was applied to a ‘reference-case’ patient based on trial data and on a large lymphoma database with approximately 1000 new patients added each year. The evaluation from the National Institute for Health and Clinical Excellence (NICE)\[30\] presented a health transition model of Markov type with five health states: PF in induction setting, PF in maintenance setting, PF not in induction/maintenance setting, PD, and death. A time horizon of 30 years was used.

3.6 Cost Data

Cost data of all evaluations are presented in table A in the SDC. In most of the studies, cost categories include expenses for chemotherapy and its administration, treatment upon relapse, routine management, and end-of-life/palliative care. Across nearly all studies, costs of rituximab and treatment costs for relapse are significant contributors to overall costs.

The presentation of cost data differs among the studies, e.g. costs of the different health states of the Markov model were described,\[34\] summarized cost parameters (for chemotherapy and follow-up) were presented,\[18\] or only chemotherapy costs and rescue therapy costs were presented.\[33\]

In none of the studies were indirect costs incorporated in the analysis.

3.7 Utilities

In five out of the 14 studies,\[16,29-31,35\] utility values were taken from a British evaluation that measured QOL in patients with FL by means of the EQ-5D instrument.\[42\] The evaluation yielded utility values of 0.805 (n = 132) for the health state PF and 0.618 for the health state PD (n = 33).

The study by Doorduijn et al.\[43\] provides the basis for utility values for several of the pharmacoeconomic studies.\[17,18,20-22,34\] In this study,\[43\] survival without events was assigned a utility of 0.83 and end-of-life care after recurrence was assigned a utility of 0.38. Utility values were measured in 119 previously untreated elderly patients (older than 65 years) with stage II-IV intermediate/high-grade NHL by means of three questionnaires: EQ-5D, EORTC-Quality of Life Questionnaire (QLQ)-C30 and the Multidimensional Fatigue Inventory (MFI-20). The model by Hornberger et al.\[32\] uses the effects of QOL in different scenarios by assigning utility weights for FL based on the study by Wild et al.\[44\] and ‘disutility’ tariffs to certain scenarios, such as chemotherapy, stem cell transplantation (SCT) and end-of-life care based on the study by Guadagnolo et al.\[45\] With concern to Wild et al.,\[44\] utility scores were obtained by using the EQ-5D instrument. In that study, 222 patients with FL were recruited. Guadagnolo et al.\[45\] did not report the method as well as the population to elicit utility values. Ray et al.\[36\] derived utility values using the EQ-5D questionnaire from a study based on a cohort of 222 patients with FL in the UK.\[46\] Ferrara and Ravasio\[33\] only carried out a cost-effectiveness evaluation with the primary endpoint of LYG.

3.8 Form of Sensitivity Analysis

While three studies reported results only from a deterministic sensitivity analysis,\[32-34\] in the remaining evaluations both deterministic and probabilistic sensitivity analyses were performed.\[16-18,20-22,29-31,35,36\]

The time horizon, cost of rituximab, changes in clinical benefit of rituximab, changes in utility values of patients with NHL, and costs of salvage therapy were the variables that were most influential in the one-way sensitivity analyses, yet the ICERs remained within acceptable ranges given the specific setting. Four studies\[18,32-34\] failed to
present a justification of the ranges of variables with concern to one-way sensitivity analysis. In all studies that had conducted a probabilistic analysis, a high probability of rituximab being a cost-effective option compared with the reference scenario was confirmed. This form of sensitivity analysis was not carried out in three evaluations.[32-34]

4. Discussion

The aim of this systematic review was to appraise the cost effectiveness of rituximab in patients with several subtypes of NHL and the methodological quality of the cost-effectiveness studies. All identified studies demonstrated that rituximab is a cost-effective option when compared with their reference scenario. However, for some studies there was uncertainty around the cost-effectiveness point estimates and the studies differed with regard to their characteristics and methodological quality. Assessing study quality, only six studies reached more than half of the attainable points.[16,17,22,30,35,36] Of remarkably high quality was the study from NICE, which reached 49 points out of 57.[30] Therefore, the conclusions drawn from these analyses should be interpreted with caution.

4.1 Methodological Rigour

On the one hand, for several dimensions of quality the majority of the studies fulfilled the specific criteria. In particular, the dimensions statement of decision problem/objective (S1), model type (S6), disease states/pathways (S8), premodel data analysis (D2), and parameter uncertainty (D4d) were fulfilled appropriately. All evaluations reported the perspective of the model clearly and provided clear definitions of the options under evaluation. Additionally, in all studies the costs and consequences that occur in the future were discounted to their present value (although studies differed with regard to the justification of the chosen discount rate).

On the other hand, several limitations that may have affected the results were identified. First, most studies did not provide a systematic method on how data were identified, nor was the quality of data assessed appropriately. Methods of elicitation of expert opinions were not described and justified. Hence, readers would be sceptical if the best available information was gathered for the analysis. To inspire a reader’s confidence in an analysis, data sources and methods for obtaining data should be presented appropriately.

Second, six studies[16,20,21,32,34,36] failed to provide sufficient justification regarding the model structure, and only five evaluations[16,17,22,30,36] explained the causal relationships described by the model structure in an appropriate way. If the structure of the model does not reflect the treatment pathways adequately, both overestimation and underestimation may occur. If information on model structure is not provided sufficiently, readers should examine whether the model structure is consistent with a coherent theory of the health condition under evaluation.

Third, only one analysis[30] considered all feasible chemotherapy options. According to the checklist by Philips et al.,[38] all clinically feasible and recommended strategies/comparators should be considered in the model. The choice of options should be described.[38] Although it would go beyond the scope of an economic evaluation to present all feasible chemotherapy options elaborately, ignoring these options makes it difficult to examine whether the applied model is best suited to the decision problem. Therefore, readers of economic evaluations should be made aware of other treatment options.

Fourth, although most of the studies failed to define and justify the cycle length in terms of the natural history of a disease, this cannot be considered as a major deficiency because the natural disease trajectory may vary. Usually, the cycle length should be the minimum interval over which the pathology or symptoms of the patient are expected to alter. By not defining and justifying the cycle length explicitly, it is difficult to interpret the results of the model comprehensively. In addition, half-cycle corrections, which adjust for the implicit bias of assuming that health state transitions are occurring at the end or the beginning of a cycle, have only been applied to three health transition models.[30,35,36]
Fifth, a probabilistic sensitivity analysis, which is recommended for the estimation of uncertainty, was not carried out in three evaluations.\textsuperscript{32-34} In addition, a considerable number of studies failed to present a justification of the ranges of variables used for the one-way sensitivity analyses.\textsuperscript{18,32-34} Plausible methods in order to define ranges for a sensitivity analysis are reviewing the literature, consulting experts or using confidence intervals around the mean for stochastic data.\textsuperscript{47}

Sixth, except for one analysis,\textsuperscript{33} in all studies a cost-utility analysis was performed. However, utility values for the calculation of QALYs are not readily available for each country and patient group. The applications of different utilities from other settings may result in different calculations and interpretations. An additional concern is that, except for five studies,\textsuperscript{18,20,22,30,36} the calculations of QALYs were not presented sufficiently.

Seventh, except for three evaluations,\textsuperscript{17,18,22} none of the economic evaluations provided ICERs for different patient groups. Additionally, only two studies\textsuperscript{32,33} discussed the generalizability of the results to other patient subgroups or settings explicitly.

Finally, internal consistency of the models in terms of their mathematical logic has only been evaluated in two evaluations.\textsuperscript{30,38} Before using them for economic analyses, models should be carefully tested in terms of accuracy and consistency.\textsuperscript{38}

One general concern is the methodological quality of the clinical studies used for data on effectiveness. With concern to our review, a pivotal trial with a Jadad score of two points or higher was regarded as methodologically sound due to practical problems in clinician blinding and ethical obligations in patient blinding (i.e. the absence of double blinding was not considered as a weakness). However, only one of the clinical trials the analyses were based on achieved a Jadad score above two points.\textsuperscript{10} Except for this one trial, the method of randomization was not addressed. Although in the sensitivity analyses of most studies varying clinical effectiveness did not affect the results significantly, an under-estimation or an overestimation of clinical effectiveness could not be ruled out.

4.2 Limitations of the Review

There were several limitations that may have affected the results of this review. First, the computerized literature search was only performed by one reviewer, while the methodological quality of the models was assessed by two reviewers. With respect to the literature search, this may draw some criticism due to selection bias. However, a detailed search strategy was performed using comprehensive searching terms in nine electronic databases.

Second, a methodological assessment could not be performed for all studies as some evaluations were only available as abstracts from professional meetings\textsuperscript{19,23-26,28} or for other reasons.\textsuperscript{27,37} The corresponding authors of the abstracts were contacted to provide more detailed information on their analyses but only one additional analysis could be obtained.\textsuperscript{16}

Finally, the authors of the checklist described several important limitations concerning its application, e.g. the checklist does not cover each aspect of a decision model.\textsuperscript{38} In addition, it does not include specific questions pertinent to particular models, and researchers cannot assess the quality appropriately without detailed knowledge about the disease.\textsuperscript{38} Therefore, the authors of the checklist recommend using it as an instrument for structuring a review. However, for the quality assessment of the studies included in this review, the checklist performed well in showing all issues concerning economic modelling that should be of particular concern to readers.

5. Conclusion and Need for Further Research

In summary, this review has shown that economic models for NHL often lack methodological quality or failed to report necessary information for readers adequately. In particular, many evaluations have shown limitations in important dimensions of quality: statement of scope/perspective (S2), rationale for structure (S3), strategies/comparators (S5), cycle length (S9), data identification (D1), QOL weights (D2d) and assessment of uncertainty (D4). Due to the prom-
isising results with regard to the clinical effectiveness of rituximab in combination with different treatment options related to different subtypes of NHL, it is expected that a number of ongoing clinical trials or updates of existing clinical trials will be published in the near future. This will result in further economic evaluations. Because the overall methodological quality of economic evaluations in this review was only intermediate, further analyses should consider all important aspects of methodological quality. Critical assessments of the methodological quality of pharmacoeconomic studies are required. The contribution of the checklist by Philips et al.\cite{38} can provide guidance in assessing the quality of pharmacoeconomic studies.

Acknowledgements

There was no funding provided for this work from any organization, company or institute. This systematic review article was written in an independent way. The authors have no financial interests to disclose directly or indirectly related to the research in the article.

All authors contributed to the study concept and data interpretation. Philipp Auweiler as the leading author is the guarantor for the overall content of the article and its interpretation. He is also responsible for the basic idea of this analysis. The authors Stephanie Stock and Andreas Gerber contributed equally to this work.

References


Correspondence: Mr Philipp Wolfgang Peter Auweiler, Clarenbachstraße 194, 50931 Cologne, Germany. E-mail: P.Auweiler@web.de