Matching-adjusted indirect comparison from the Lymphoma Epidemiology of Outcomes Consortium for Real World Evidence (LEO CReWE) study to a clinical trial of mosunetuzumab in relapsed or refractory follicular lymphoma

Matthew J. Maurer,[1,2] Carla Casulo,[3] Melissa C. Larson,[1] Thomas M. Habermann,[2] Izidore S. Lossos,[4] Yucai Wang,[2] Loretta J. Nastoupil,[5] Christopher Strouse,[6] Dai Chihara,[5] Peter Martin,[7] Jonathon B. Cohen,[8] Brad S. Kahl,[9] W. Richard Burack,[3] Jean L. Koff,[8] Yong Mun,[10] Anthony Masaquel,[10] Mei Wu,[10] Michael C. Wei,[10] Ashwini Shewade,[10] Jia Li,[10] James R. Cerhan,[1] Brian K. Link[6] and Christopher R. Flowers[5]

[1]Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN; [2]Division of Hematology, Mayo Clinic, Rochester, MN; [3]Department of Medicine, Wilmot Cancer Institute, University of Rochester, Rochester, NY; [4]Department of Medicine, Comprehensive Sylvester Cancer Center, University of Miami, Miami, FL; [5]University of Texas, MD Anderson Cancer Center, Houston, TX; [6]Department of Medicine, University of Iowa, Iowa City, IA; [7]Department of Medicine, Weill Cornell Medical College, New York, NY; [8]Department of Hematology and Medical Oncology, Winship Cancer Institute, Emory University, Atlanta, GA; [9]Department of Medicine, Washington University in St Louis, St Louis, MO and [10]Genentech, Inc., South San Francisco, CA, USA

Correspondence: M.J. Maurer
maurer.matthew@mayo.edu

Received: June 14, 2023.
Accepted: November 17, 2023.
Early view: November 30, 2023.

https://doi.org/10.3324/haematol.2023.283737

Abstract

Mosunetuzumab is a novel bispecific antibody targeting epitopes on CD3 on T cells and CD20 on B cells with the goal of inducing T-cell mediated elimination of malignant B cells. A recent pivotal phase I/II clinical trial (GO29781) demonstrated that mosunetuzumab induced an overall response rate (ORR) of 80%, complete response (CR) rate of 60%, and a median progression-free survival (PFS) of 17.9 months in patients with relapsed/refractory (R/R) follicular lymphoma (FL) following at least two prior lines of systemic therapy, including alkylator and anti-CD20 antibody-based therapy. Historical data from cohorts receiving therapy for R/R FL can provide some context for interpretation of single-arm trials. We compared the results from the mosunetuzumab trial to outcomes from a cohort of patients with R/R FL from the LEO Consortium for Real World Evidence (LEO CReWE). We applied clinical trial eligibility criteria to the LEO CReWE cohort and utilized matching-adjusted indirect comparison weighting to balance the clinical characteristics of the LEO CReWE cohort with those from the mosunetuzumab trial. ORR (73%, 95% CI: 65-80%) and CR rates (53%, 95% CI: 45-61%) observed in the weighted LEO CReWE cohort were lower than those reported on the mosunetuzumab trial (ORR=80%, 95% CI: 70-88%; CR=60%, 95% CI: 49-70%, respectively). PFS at 12 months was similar in the weighted LEO CReWE (60%, 95% CI: 51-69%) and the mosunetuzumab (58%, 95% CI: 47-68%) trial. Sensitivity analyses examining the impact of matching variables, selection of line of therapy, and application of eligibility criteria provide context for best practices in this setting.

Introduction

Follicular lymphoma (FL) is a highly heterogeneous disease often characterized as indolent in behavior requiring intermittent systemic therapy over time.[1] Although most patients with FL will experience a life expectancy comparable to that of the general population, a subset will have early disease-related mortality, often preceded by early relapse following initial immunochemotherapy (IC), refractoriness to alkylator therapies, or transformation to aggressive lymphoma.[2] Patients with relapsed or refractory (R/R) FL have several therapeutic options available without

an agreed-upon standard.[3] As such, selection of therapy in the third line or later requires thoughtful consideration of patient characteristics, prior therapy, expected toxicity, and disease behavior. Several reports of longitudinal FL patient cohorts illustrate that, on average, patients experience progressively shorter response durations or treatment-free intervals over sequential treatment lines.[4-7] Thus, patients needing treatment after two or more prior therapies (especially if refractory to prior agents) represent a population with unmet needs that may be addressed by novel strategies.

Mosunetuzumab is a novel T-cell engaging bispecific antibody targeting epitopes on CD3 on T cells and CD20 on B cells with the goal of inducing T-cell-mediated elimination of malignant B cells. A recent pivotal phase I/II clinical trial (GO29781; clinicaltrials.gov identifier: NCT02500407) demonstrated that mosunetuzumab induced an overall response rate (ORR) of 80%, a complete response (CR) rate of 60%, and a median progression-free survival (PFS) of 17.9 months in patients with R/R FL following at least two prior lines of systemic therapy, including alkylator and anti-CD20 antibody-based therapy.[8] Single-arm clinical trial designs such as GO29781 result in limited ability to determine how well unmet clinical needs have been addressed due to lack of a control cohort. Historical data from cohorts receiving therapy for R/R FL can provide some context for interpretation,[4,9] but may suffer from a composition of patients that differ substantially from the trial cohort of interest in any of several relevant characteristics such as prior treatment histories, drug class refractoriness, and evolving patterns of care. We have recently demonstrated that, in expert academic US practices, treatment selection for R/R FL is quite variable and reflects disease heterogeneity.[9] The aims of the present study were to provide a focused comparison of outcomes in a comparator real-world population to the GO29781 trial and evaluate potential best practices for implementing these types of comparison studies in the setting of R/R FL.

Methods

The Lymphoma Epidemiology of Outcomes Consortium for Real World Evidence (LEO CReWE) was used to build a real-world evidence cohort of patients with R/R FL who received at least two prior lines of systemic therapy; full details can be found in our previous publication.[9] The study was approved by the Mayo Clinic Institutional Review Board. The population for the primary analysis comprised patients who met all of the following eligibility criteria: i) received systemic therapy for FL grade 1-3A after at least two prior lines of systemic therapy that included an anti-CD20-directed therapy and an alkylating agent; ii) met all key eligibility criteria from the GO29781 trial (listed in *Online Supplementary Appendix*) for a potential index line of therapy with no missing eligibility or matching data; and iii) did not have transformed disease prior to a potential index line of therapy (Figure 1). Matching variables were as follows: age (years) at index therapy (mean, Standard Deviation [SD]); prior lines of therapy (mean, SD); progression of disease within 24 months (POD24) following front-line IC (yes *vs*. no *vs*. did not receive IC); double-refractory to anti-CD20 and alkylator therapy (yes *vs*. no); and elevated lactate dehydrogenase (LDH) at index therapy (yes *vs*. no). Matching-adjusted indirect comparison[10] (MAIC) weighting was performed on the LEO CReWE dataset to select the index line for each patient and generate weights for comparison to the GO29781 study. The primary outcome measure for this study was ORR, defined as the proportion of patients with best response as CR or partial response (PR) during the available follow-up beginning from the start of therapy (index line) to the earliest of the following: documented progression of disease, initiation of a new line of anti-lymphoma therapy, transformation to an aggressive lymphoma, death, or end of follow-up as per the pre-specified clinical cut-off date. Secondary outcome measures were CR rate and PFS. Further details are provided in the *Online Supplementary Appendix*.

Statistical analysis

The statistical analysis plan (SAP) was jointly developed and finalized prior to the primary analysis. All analyses were performed by members of the LEO Cohort Statistics and Informatics Core (MJM, MCL). MAIC,[10] a form of propensity score weighting, was applied to individual patient data (IPD) from the LEO CReWE study.[11] A series of sensitivity analyses was performed to evaluate the impact of study outcomes when utilizing alternate approaches to the following: i) inclusion/exclusion criteria; ii) choice of matching variables; and iii) method for selecting which line of therapy to use for a given patient when multiple lines of therapy meet study eligibility. Continuous variables were summarized using descriptive statistics (median, interquartile range [IQR], or mean, SD); categorical variables, including response rates, were summarized as proportions and/or rates. Time-to-event variables were summarized using Kaplan-Meier curves and 95% Confidence Intervals (CI). Associations between groups and categorical endpoints (e.g., ORR and CR) were assessed using logistic regression and summarized with Odds Ratios and 95% CI. Further details on the study analyses can be found in the SAP (*Online Supplementary Appendix*).

Results

LEO CReWE cohort

The starting population for the study consisted of 441 patients who received systemic therapy for FL grade 1-3A following at least two prior lines of systemic therapy that

included an anti-CD20-directed therapy and an alkylating agent. Seventy-three patients (17%) were excluded due to presence of variables that did not meet the inclusion / exclusion criteria based on the GO29781 clinical trial (*Online Supplementary Table S1*), and 157 patients (36%) were excluded due to missing data on one or more variables from inclusion / exclusion criteria and inability to confirm all GO29781 eligibility criteria, yielding 211 patients for the application of the MAIC analysis. An additional 9 patients were missing one or more of the matching variables, resulting in 202 LEO CReWE patients for the primary analysis. Full details on the 202 LEO CreWE patients prior to weighting can be found in *Online Supplementary Table S2*. The index therapy utilized in the primary analysis was 3rd line for 116 patients (57%), 4th line for 48 patients (24%), and 5th line or later for 38 patients (19%). Median age at index therapy was 60 years (IQR: 53-68), and 118 patients (58%) were male. Most patients (N=157, 78%) had stage III/IV disease at index therapy; 59 patients (29%) had elevated LDH at index therapy and 31 (15%) had bulky disease (>7 cm). Eighty-six patients (43%) experienced POD24 following first-line IC, 63 patients (31%) had disease progression after 24 months to front-line IC, and 53 patients (26%) did not receive first-line IC. One hundred and twenty-three patients (61%) had received prior anthracycline, and 26 patients (13%) had received a prior autologous stem cell transplant. One hundred and forty-one patients (70%) had FL refractory to previous CD20 antibody therapy, with 38% refractory to alkylating agents, and 69% refractory to their most recent prior line of therapy. The median time from diagnosis to index line was 62 months (IQR 41-93), and the median time from start of prior therapy to start of index line was 11 months (IQR 4-15). The most common class of therapy at index line was CD20 antibody-based IC (N=57, 28%). Additional therapy class included novel therapies with or without CD20 antibody (N=47, 23%), CD20 antibody and lenalidomide-based therapy (N=28, 14%), platinum-based salvage chemotherapy (N=23, 11%), and CD20 antibody monotherapy (N=18, 9%) (*Online Supplementary Table S3*). Twenty-eight patients (14%) received a stem cell transplant as index therapy (N=18 autologous and N=10 allogeneic), 7 received chimeric antigen receptor (CAR) T-cell therapy, and 4 received a bi-specific antibody; 80 patients (40%) received index therapy on a clinical trial (*Online Supplementary Table S4*). Most patients (N=171, 85%) received their index therapy at one of the eight LEO institutions. Response assessment on index therapy was positron emission tomography (PET)- or PET/computed tomography (CT)-based in 109 patients (54%), CT-based in 73 patients (36%), and clinical or unspecified in 20 patients (10%). At a median follow-up of 58.0 months from start of index therapy (range 0.03-252), 107 patients (53%) had had a PFS event after index therapy, and 47 patients (23%) had died. The unweighted CR and ORR in the 202 patients were 58% (95% CI: 51-65) and 78% (95% CI: 71-83), respectively; the unweighted 12-month PFS was 65% (95% CI: 59-72).

GO29781 trial

Details on the GO29781 trial have been previously reported.[8] Most relevant to this analysis, median age at study entry in the 90 patients enrolled was 60 years (IQR 53-67), and 61% were male. LDH was elevated in 39% at study entry. Median number of previous lines of therapy was 3 (IQR 2-4), with 38% receiving 2 previous lines, 31% receiving 3 previous lines, and 31% receiving more than 3 previous lines. Thirty-eight patients (42%) had POD24 to front-line immunochemotherapy, and 48 (53%) had double-refractory disease to previous anti-CD20 and alkylator therapies.

Primary comparison of LEO CReWE to GO29781 by matching-adjusted indirect comparison weighted analysis

Prior to weighting, significant differences were observed in the clinical characteristics of the primary analysis cohort selected from the LEO CReWE (N=202) and the GO29781 trial (Table 1). LEO CReWE patients were less heavily pre-treated and had fewer cases of double refractory disease. LEO CReWE patients also had a longer time from previous treatment to their index therapy. Application of the MAIC utilizing the 5 matching variables resulted in an effective sample size (ESS; the number of independent non-weighted individuals that would be required to give an estimate with the same precision as the weighted sample estimate) of 127.3 and a weighted N of 167; the distribution of the weights is shown in *Online Supplementary Figure S1*. MAIC weighting rebalanced the distributions of all 5 matching variables, including number of prior lines of therapy. Other key clinical variables were also better aligned between the LEO CReWE and GO29781 studies after application of weights (Table 1). ORR observed in the weighted LEO CReWE cohort (73%, 95% CI: 65-80%) was lower than the ORR (80%, 95% CI: 70-88%) reported on the GO29781 trial (Table 2). Similarly, CR in the weighted LEO CReWE cohort (53%, 95% CI: 45-61%) was lower than the GO29781 trial (60%, 95% CI: 49-70%). PFS at 12 months was similar in the LEO CReWE cohort (12-month PFS 60%, 95% CI: 51-69%) and the GO29781 trial (12-month PFS 58%, 95% CI: 47-68%) (Figure 2).

Sensitivity analyses

We performed a series of pre-specified sensitivity analyses to determine the impact of inclusion/exclusion criteria, matching variables, and index therapy selection on the outcomes in the study (Table 3). Application of GO29781-based trial criteria and the requirement of non-missing data in the LEO CReWE cohort resulted in exclusion of over half of potential patients (230 out of 441 potential patients, 52%) from the primary analysis. Utilizing an alternative set of clinical trial inclusion / exclusion criteria had little impact on sample size or study results compared to the primary analysis. Ignoring the clinical trial eligibility criteria would have greatly increased the overall (N=357) and ef-

Table 1. Comparison of key patient characteristics in GO29781 *versus* LEO CReWE cohort (unweighted and matching-adjusted indirect comparison weighted).

Variable	GO29781 N=90	LEO CReWE (unweighted) N=202	Delta	Delta *P*	LEO CReWE (MAIC weighted) Weighted N=167 ESS=127	Delta	Delta *P*
Used in MAIC matching							
Age in years, mean (SD)	60.0 (12.0)	60.2 (10.8)	0.2	0.85	60.3 (10.5)	0.3	0.69
Elevated LDH, %	39	29	-10	0.13	39	0	1.00
POD24 to 1L IC, %	42	43	1%	1.00	42	0	1.00
Prior LOT, mean (SD)	3.3 (1.7)	2.7 (1.1)	-0.57	<0.001	3.3 (1.8)	-0.03	0.83
Double refractory, %	53	36	-17	0.009	53	0	1.00
Not used in MAIC matching							
Male, %	61	58	-3	0.76	58	-3	0.68
Bulky disease, %	18	16	-2	0.73	12	-6	0.32
Stage III/IV, %	77	84	7	0.96	80	4	0.62
Prior SCT, %	21	13	-8	0.10	15	-6	0.32
Months since prior therapy, mean (SD)	14.2 (16.9)	18.6 (21.1)	4.4	0.004	14.8 (19.2)	0.6	0.68

N: number; MAIC: matching-adjusted indirect comparison; ESS: effective sample size; LDH: lactate dehydrogenase; POD24: progression of disease in 24 months; 1L: first-line; IC: immunochemotherapy; LOT: line of therapy; SCT: stem cell transplant.

Table 2. Primary results: comparison of GO29781 to LEO CReWE cohort.

Group	N (evaluable for response)	ORR (95% CI)	CR rate (95% CI)	PFS12 (95% CI)
LEO CReWE (unweighted)	202 (192)	77.6 (70.9-83.2)	57.8 (50.5-64.8)	65.0 (58.6-72.2)
LEO CReWE (MAIC weighted)	167 (160)	73.0 (65.3-79.5)	52.9 (44.8-60.7)	59.5 (51.0-69.3)
GO29781 (trial results)	90 (90)	80.0 (70.3-87.7)	60.0 (49.1-70.2)	57.7 (46.9-68.4)

N: number; ORR: overall response rate; CI: confidence interval; CR: complete response; PFS12: progression-free survival at 12 months.

fective (N=297) sample size of the study but made only a small impact on weighted estimates (ORR=74%, CR=56%, 12-month PFS 63%). Utilizing different sets of matching variables and/or variable types (e.g., dichotomous *vs*. continuous) in the MAIC implementation also had little effect on the ORR, CR, and PFS estimates. The greatest impact was seen when changing the algorithm for selecting the index line of therapy for patients with multiple eligible lines of therapy. Using a randomly selected treatment line in the LEO CReWE Cohort yielded a lower ESS (N=84.7) (*Online Supplementary Figure S2*), though clinical characteristics remained well balanced after weighting. Outcomes in the LEO CReWE cohort were also inferior when using a random line (ORR=72%, CR=42%, 12-month PFS=54%) compared to the primary analysis. Utilizing the first eligible line for a patient resulted in an essentially unusable analysis, with an ESS of 20.1 and most patients having near-zero weights (*Online Supplementary Figure S3*). Selecting the last eligible line yielded nearly identical results to the primary analysis, due to the preferential weighting of later lines of therapy in LEO CReWE. Summary details of the sensitivity analysis results can be found in Table 3.

In a post-hoc subset analysis of LEO CReWE by receipt of index therapy on clinical trial, response rates (ORR=76%, 95% CI: 65-85%; CR=56%, 95% CI: 44-67%) were higher and PFS12 was lower (57%, 95% CI: 44-73%) on trials compared to index therapy received off-trial (ORR=70%, 95% CI: 59-79%; CR=50%, 95% CI: 40-61%; PFS12=62%, 95% CI: 51-75%) (*Online Supplementary Table S5*).

Discussion

In this study, we evaluated response rates and PFS in an observational cohort of patients with FL treated in the third line or later after prior alkylator and anti-CD20-based therapy. The LEO CReWE cohort was restricted to those meeting key eligibility criteria in the GO29781 trial and further re-weighted to align with the clinical characteristics of the patients treated on GO29781. Strengths of the study include the assembly of a large observational cohort of patients with R/R FL and detailed clinical annotation and outcomes from diagnosis through all lines of therapy. LEO CreWE patients with R/R FL were treated in medical cen-

ters and clinical contexts similar to those for the GO29781 trial, with 40% of patients receiving index therapy on a clinical trial, which is distinct from observational studies involving cohorts of patients with R/R FL from general practices.[12-14] MAIC rebalanced the LEO CReWE cohort to align with the clinical characteristics of patients enrolled on the GO29781 trial. Standard limitations of these types of comparative effectiveness studies apply, including potential bias by both observed and unobserved differences in the patients between the two cohorts. Adverse event data were not available in the LEO CReWE cohort, and thus this study does not provide clinical comparisons of toxicity or tolerability. Safety profiles and quality of life should be considered when making clinical decisions regarding therapy in this clinical space.

Individual patient-level data (IPD) from GO29781 were not available for this analysis, and thus the weighting of the LEO CReWE data was based on the summary statistics for the GO29781 trial data as opposed to matching or propensity weighting approaches utilizing IPD. The MAIC approach utilized in this study weighted the individuals in the LEO CReWE cohort such that the key clinical characteristics used for matching were balanced between the LEO CReWE and GO29781. Using this approach, all eligible patients from the LEO CReWE cohort were utilized in the analysis, but LEO CReWE patients with disparate clinical characteristics compared to the GO29781 trial participants contributed less information to the analysis compared to LEO CReWE patients with similar clinical characteristics as the GO29781 trial. Methods for checking fit and calibration in a MAIC weighting approach are limited compared to traditional propensity score weighting, in which IPD is available. Additionally, residual confounding may remain after MAIC weighting.

Response rates observed in the LEO CReWE cohort (weighted ORR=73%, CR=53%) were encouraging for R/R FL, yet remained lower than those observed in the GO29781 study (ORR=80%, CR=60%). PFS at 12 months was similar between the LEO CReWE cohort and GO29781 trial based on the reported duration of follow-up used in this analysis. Notably, response rates (ORR=76%, CR=56%) were higher and PFS12 (57%) was lower for LEO CReWE index therapies received on clinical trial, though this may be confounded by patient selection and varying response rates of different treatment classes. Response and progression are assessed differently between clinical trials and routine clinical practice, which needs to be considered in the interpretation of these results. Clinical trials feature frequent and standardized disease assessment that includes imaging at regular intervals. In routine clinical care, intervals for response assessment are not stringently dictated, and imaging may not necessarily be performed on a regular schedule. Further, clinicians may not universally confirm response with repeat bone marrow biopsy (BMB) in patients who had prior bone marrow involvement, and clinical plans to proceed to

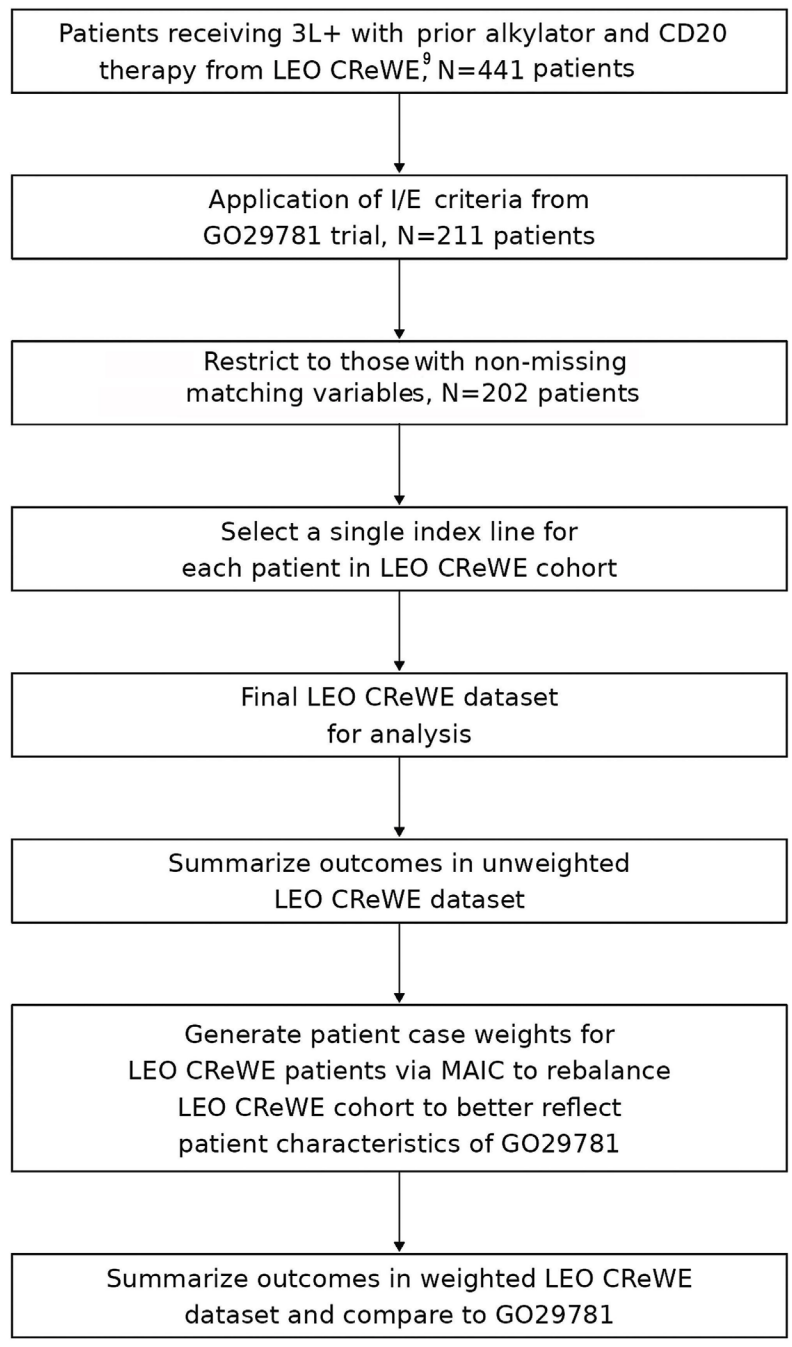

Figure 1. Analysis flow chart. I / E: inclusion / exclusion; MAIC: matching-adjusted indirect comparison; N: number; PFS: progression-free survival.

transplantation based on response may influence decision making. In the GO29781 trial, a BMB was required at study entry with repeat BMB to confirm complete response if bone marrow involvement was present at baseline. This was not required in the LEO CReWE cohort and only 41% of patients had bone marrow biopsied at the start of index therapy (*data not shown*). However, in an analysis of 580 patients pooled from 7 National Clinical Trials Network (NCTN) clinical trials of FL, Rutherford and colleagues identified that BMB were irrelevant to assessing complete response in 99% of patients.[15] Thus it is unlikely that differences in routine BMB between trial and observational cohorts has a significant impact on the interpretation of results. Unlike aggressive lymphoma, where progression often manifests clinically prior to planned imaging,[16] routine imaging on clinical trials may contribute to shorter PFS in the setting of FL (as compared to observational studies without routine imaging) when an increase in tumor size may not correspond to a clinical indication for evaluation or subsequent therapy. A prior study from three centers

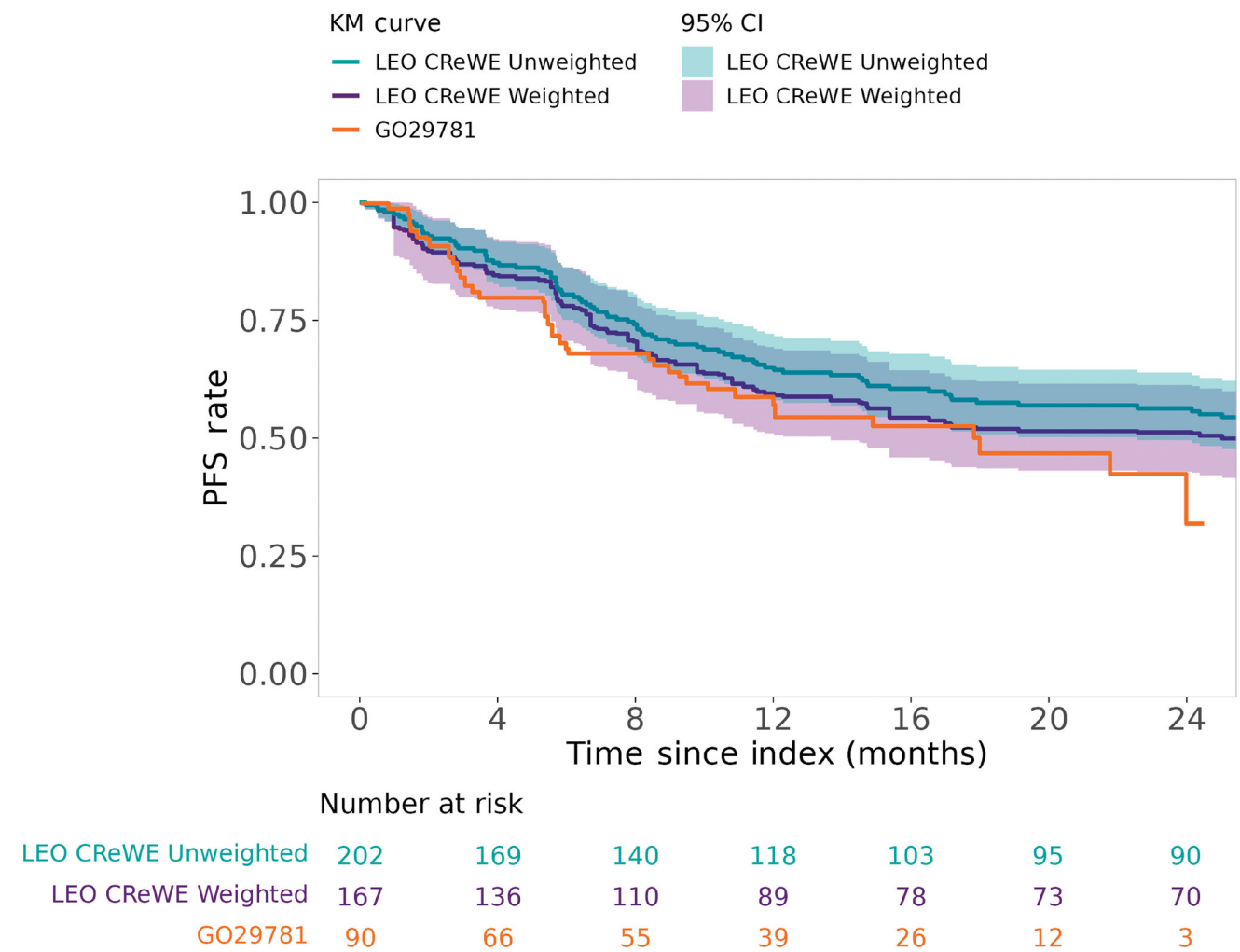

Figure 2. Kaplan-Meier curve of progression-free survival for LEO CReWE *versus* GO29781 trial. CI: Confidence Interval; N: number; PFS: progression-free survival.

within LEO CReWE identified that 54-64% of relapses to front-line therapy for FL were detected clinically.[17] A distinct advantage of the LEO consortium is the involvement of experts in lymphoma clinical trials and patient management, including leaders in the NCTN and several members of the NCI Lymphoma Steering Committee, who have expertise in lymphoma response assessment.

Despite these commonly cited limitations for synthetic cohorts, several recent studies[18-20] have used this strategy to support the impact of phase II trials in lymphoma. Our study differs notably from the others by demonstrating modest differences between the measured outcomes of the phase II trial results and the synthetic cohort. There are several possible explanations for this difference, including the unique choice of MAIC methodology to accommodate lack of IPD from GO29781, the line of index therapy or variables chosen for matching, improved assessment of outcomes and/or improved outcomes in an observational cohort managed by lymphoma experts, or the true relative activity of mosunetuzumab compared to other options in this patient setting.

The generation of a synthetic cohort and/or implementation of a matching-based analysis requires a series of decisions that may influence the results. Analytical decisions such as eligibility criteria, selection of index therapy, and matching variables are then applied to the cohort selected for the analysis. We performed a series of sensitivity analyses to address the uncertainty in estimation of outcomes introduced by such decisions. Selecting the cohort and index line is especially critical given the heterogeneity in treatment selection and expected outcomes for patients with R/R FL. The index line of therapy for LEO CReWE was chosen based on the line of therapy most like the clinical characteristics of the GO29781 trial based on MAIC weights. Our initial publication of the LEO CReWE FL cohort primarily focused on outcomes in patients at initial time of eligibility for the GO29781 trial (i.e., third line or beyond with prior anti-CD20 and alkylator therapy), with 94% of patients achieving this in the third line setting.[9] However, utilizing the first eligible line in our sensitivity analysis resulted in essentially unusable results due largely to imbalances in the number of prior lines of therapy. This highlights the importance of careful examination of clinical characteristics and thoughtful alignment of a synthetic cohort with clinical trial criteria. In contrast, alternative collections of matching variables and trial inclusion / exclusion criteria had little impact on effective sample size or weighted outcomes.

Conclusions

These results support our previously published data showing that, despite multiple recurrences, patients

Table 3. Sensitivity analyses of matching-adjusted indirect comparison weighting scenarios in LEO CReWE cohort.

	Scenario	N	ESS	MAIC weighted ORR (95% CI)	MAIC weighted CR (95% CI)	MAIC weighted PFS12 (95% CI)
	Unweighted analysis	202	-	77.6 (70.9-83.2)	57.8 (50.5-64.8)	65.0 (58.6-72.2)
	Primary MAIC analysis	202	127	73.0 (65.3-79.5)	52.9 (44.8-60.7)	59.5 (51.0-69.3)
Change trial I/E criteria application						
1	No trial I/E criteria	357	297	73.7 (68.2-78.6)	56.0 (50.1-61.8)	63.1 (57.3-69.5)
2	LEO clinician trial I/E criteria	217	172	74.1 (67.0-80.2)	54.6 (47.1-61.9)	64.7 (57.3-73.0)
Change matching variables						
3	Add gender	202	128	72.9 (65.2-79.5)	52.3 (44.3-60.2)	59.2 (50.8-69.1)
4	Add stage	190	116	70.6 (62.4-77.7)	52.0 (43.6-60.3)	59.5 (50.7-69.8)
5	Add ECOG PS	183	114	73.6 (65.5-80.4)	52.7 (44.2-61.0)	62.1 (53.2-72.5)
6	Add FLIPI	188	117	69.3 (61.0-76.5)	49.1 (40.8-57.4)	56.5 (47.8-66.8)
7	Add bulky disease	191	120	70.4 (62.4-77.4)	52.8 (44.6-60.8)	55.6 (47.0-65.9)
8	Substitute refractory to prior line	203	132	71.7 (64.1-78.4)	51.1 (43.2-59.0)	58.4 (50.1-68.1)
9	Substitute POD24 to any 1L	201	127	74.4 (66.9-80.8)	53.3 (45.3-61.1)	59.9 (51.4-69.7)
10	Dichotomize prior LOT	202	133	66.7 (58.7-73.9)	50.6 (42.6-58.6)	55.3 (47.0-64.9)
Change selection of index therapy						
11	Random line	202	85	71.5 (63.1-78.8)	41.7 (33.4-50.5)	54.1 (43.8-66.8)
12	First eligible line	202	20	49.4 (36.5-62.5)	34.8 (23.0-48.5)	34.4 (21.2-55.8)
13	Last eligible line	202	126	73.7 (66.1-80.2)	53.1 (45.1-61.0)	61.5 (53.1-71.3)

N: number; MAIC: matching-adjusted indirect comparison; ESS: effective sample size; CI: confidence interval; POD24: progression of disease in 24 months; 1L: first-line; LOT: line of therapy; ORR: overall response rate; CR: complete response; PFS12: progression-free survival at 12 months; I/E: inclusion / exclusion; ECOG PS: Eastern Cooperative Oncology Group Performance Status; FLIPI: Follicular Lymphoma International Prognostic Index.

with R/R FL respond favorably to therapy, albeit of limited duration. These data suggest that the encouraging response rates observed in this novel class of bispecific therapy in a heavily pre-treated population yield similar PFS to our comparison cohort with the current study follow-up. Although the methodology utilized in this analysis has limitations relative to a randomized clinical trial, it helps to provide comparative context for clinical outcomes and patterns of care. Differences in response and progression assessment methodology should be taken into consideration when making direct comparisons between clinical trials and observational cohorts of patients treated in routine clinical practice. Development of a set of best practices by clinical expert consensus for these types of comparative effectiveness analyses in the R/R FL space may be beneficial for more consistency in future studies. Comprehensive data that include safety, tolerability, quality of life, as well as efficacy, should be considered when evaluating treatment options for patients with R/R FL.

Disclosures

MJM declares advisory board (Adaptive Biotechnologies, AstraZeneca, GenMab); consulting (BMS); research funding (Genmab, Genentech/Roche, BMS); spouse employment and stock (Exact Sciences). CC declares research funding (SecuraBio, Genmab, Gilead, Genentech, Inc.). MCL declares research funding (Genentech, Genmab). TMH declares research funding (Genentech, Sorrento); consulting or advisory role (Eli Lilly & Co., Morphosys, Incyte, Biegene, Loxo Oncology); data monitoring committee (Seagen, Tess Therapeutics, Eli Lilly & Co.). ISL declares current employment (University of Miami); honoraria (Adaptive); consulting or advisory role (LRF); research funding (NCI); provided compensation teaching (Kyowa Kirin). YW declares research funding (Incyte, InnoCare, LOXO Oncology, Eli Lilly, MorphoSys, Novartis, Genentech, Genmab); advisory board (Eli Lilly, LOXO Oncology, TG Therapeutics, Incyte, InnoCare, Kite, Jansen, BeiGene); consultancy (Innocare, AbbVie); honorarium (Kite). LJN declares research funding (BMS, Caribou Biosciences, Epizyme, Genentech, Gilead/Kite, Genmab, Janssen, IGM Biosciences, Novartis, Takeda); honoraria (Abbvie, ADC Therapeutics, BMS, Caribou Biosciences, Epizyme, Genentech/Roche, Gilead/Kite, Genmab, Incyte, Janssen, MEI, Novartis, Takeda); DSMC (Genentech/Roche, MEI, Takeda, Denovo). CS declares current employment (University of Iowa); honoraria (Iowa Oncology Society, Curio Science); consulting or advisory role (Pfizer, GSK). DC declares research funding (Janssen, Epizyme, BMS, MorphoSys). PM declares consulting or advisory role (Abbvie, AztraZeneca, Beigene, BMS, Daiichi, Genentech, Janssen). JBC

declares consulting or advisory role (BeiGene, ADCT, Lilly, AstraZeneca, Janssen, Genentech, Inc., HutchMed, Kite/Gilead); research funding (Genentech, Inc., Celgene, Novartis, Takeda, AstraZeneca, Loxo/Lilly, Lam Therapuetics, BioInvent); Expert Testimony (Hawaii). BSK declares consultancy (Astrazeneca, ADCT Therapeutics, Roche, Genetech, Abbvie, MEI, AcertaPharma,Pharmacyclics, Celgene/BMS, Beigene, Kite, Janssen, Incyte, Hutchmed, TG Therapeutics, BeiGene, Genmab, Seattle Genetics); research funding (Genetech, Abbvie, Celgene/BMS, BeiGene Astra Zeneca, Hutchmed); speaker's bureau (Research to Practice). JLK declares consulting or advisory role (Gamida Cell, MorphoSys, BeiGene USA, TG Therapeutics); research funding (Viracta Therapeutics). YM declares current employment (Genentech/Roche); stock ownership (Genentech/Roche). AM declares current holder of individual stocks in a privately-held company (Roche); current holder of stock options in a privately-held company (Roche). MCW declares current employment (Genentech); stock ownership (Roche); travel, accommodation, and expenses (Genentech/Roche). AS declares current employment (F. Hoffmann-La Roche Ltd.); stock ownership (F. Hoffmann-La Roche Ltd). JL declares stock ownership (F. Hoffmann-La Roche Ltd.). JC declares consulting or advisory role (BMS, Protagonist Therapeutics SMC). BKL declares consulting or advisory role (Genentech/Roche, MEI Pharma Inc.); research funding (Genentich/Roche, Genmab, Celgene, AstraZeneca); expert testimony (Amgen). CRF declares consulting or advisory role (Bayer, Gilead, Spectrum Pharmaceuticals, AbbVie, Celgene, Denovo Biopharma, BeiGene, Karyopharm Therapeutics, Pharmacyclics/Janssen, Genentech/Roche, Epizyme, Genmab, Seattle Genetics, Foresight Diagnostics, BMS, Curio Science, AstraZeneca, MorphoSys); stock ownership (Foresight Diagnostics, N Power); research funding (Acerta Pharma, Janssen, Gilead, Celgene, TG Therapeutics, Genentech/Roche, Pharmacyclics, AbbVie, Millennium, Alimera Sciences, Xencor, 4D Pharma, Adaptimmune, Amgen, Bayer, Cellectis, EMD Serono, Guardant Health, Iovance Biotherapeutics, Kite/Gilead, MorphoSys, Nektar, Novartis, Pfizer, Sanofi, Takeda, Ziopharm Oncology). WRB has no conflicts of interest to disclose.

Contributions

MM, CC, ML, JL, TH, IL, YW, LN, CS, DC, PM, JC, BK, WB, JK, YM, AM, MW, MCW, AS, JL, JC, BL and CF are responsible for the conception and/or design of the study/research, for analysis and/or interpretation of the data, and wrote the manuscript. MM, CC, ML, JL, TH, IL, YW, LN, CS, DC, PM, JC, BK, WB, JK, JC, BL and CF are responsible for the collection and/or assembly of the data.

Funding

This work was supported by a grant from the National Cancer Institute: Lymphoma Epidemiology of Outcomes (U01CA195568) and a research grant from Genentech. No medical writer was used for this manuscript.

Data-sharing statement

The data in the study are not publicly available. Data sharing policies and the process to request the LEO CReWE data that support the findings of this study can be found on the LEO Cohort website: https://leocohort.org/

References

1. Jacobsen E. Follicular lymphoma: 2023 update on diagnosis and management. Am J Hematol. 2022;97(12):1638-1651.
2. Sarkozy C, Maurer MJ, Link BK, et al. Cause of death in follicular lymphoma in the first decade of the rituximab era: a pooled analysis of French and US cohorts. J Clin Oncol. 2019;37(2):144-152.
3. Qualls D, Salles G. Prospects in the management of patients with follicular lymphoma beyond first-line therapy. Haematologica. 2022;107(1):19-34.
4. Link BK, Day BM, Zhou X, et al. Second-line and subsequent therapy and outcomes for follicular lymphoma in the United States: data from the observational National LymphoCare Study. Br J Haematol. 2019;184(4):660-663.
5. Rivas-Delgado A, Magnano L, Moreno-Velázquez M, et al. Response duration and survival shorten after each relapse in patients with follicular lymphoma treated in the rituximab era. Br J Haematol. 2019;184(5):753-759.
6. Batlevi CL, Sha F, Alperovich A, et al. Follicular lymphoma in the modern era: survival, treatment outcomes, and identification of high-risk subgroups. Blood Cancer J. 2020;10(7):74.
7. Salles G, Schuster SJ, Fischer L, et al. A retrospective cohort study of treatment outcomes of adult patients with relapsed or refractory follicular lymphoma (ReCORD-FL). Hemasphere. 2022;6(7):e745.
8. Budde LE, Sehn LH, Matasar M, et al. Safety and efficacy of mosunetuzumab, a bispecific antibody, in patients with relapsed or refractory follicular lymphoma: a single-arm, multicentre, phase 2 study. Lancet Oncol. 2022;23(8):1055-1065.
9. Casulo C, Larson MC, Lunde JJ, et al. Treatment patterns and outcomes of patients with relapsed or refractory follicular lymphoma receiving three or more lines of systemic therapy (LEO CReWE): a multicentre cohort study. Lancet Haematol. 2022;9(4):e289-e300.
10. Signorovitch JE, Sikirica V, Erder MH, et al. Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research. Value Health. 2012;15(6):940-947.
11. Phillippo DM, Ades AE, Dias S, Palmer S, Abrams KR, Welton NJ. Methods for population-adjusted indirect comparisons in health technology appraisal. Med Decis Making. 2018;38(2):200-211.
12. Albarmawi H, Nagarajan M, Onukwugha E, et al. Follicular lymphoma treatment patterns between 2000 and 2014: a SEER-Medicare analysis of elderly patients. Future Oncol. 2020;16(8):353-365.
13. Xie C, Li R, Huang X, Chihara D, Flowers CR. The impact of

sequence of therapy for older patients with follicular lymphoma: SEER-Medicare analysis. Clin Lymphoma Myeloma Leuk. 2022;22(10):e938-e946.
14. Ayers EC, Margolis D, Landsburg DJ. Real world outcomes in patients with relapsed/refractory diffuse large B-cell lymphoma receiving palliative intent therapies. Clin Lymphoma Myeloma Leuk. 2020;20(10):661-667.
15. Rutherford SC, Yin J, Pederson L, et al. Relevance of bone marrow biopsies for response assessment in US National Cancer Institute National Clinical Trials Network Follicular Lymphoma Clinical Trials. J Clin Oncol. 2023;41(2):336-342.
16. Thompson CA, Ghesquieres H, Maurer MJ, et al. Utility of routine post-therapy surveillance imaging in diffuse large B-cell lymphoma. J Clin Oncol. 2014;32(31):3506-3512.
17. Goldman ML, Mao JJ, Strouse CS, et al. Surveillance imaging during first remission in follicular lymphoma does not impact overall survival. Cancer. 2021;127(18):3390-3402.
18. Neelapu SS, Locke FL, Bartlett NL, et al. Comparison of 2-year outcomes with CAR T cells (ZUMA-1) vs salvage chemotherapy in refractory large B-cell lymphoma. Blood Adv. 2021;5(20):4149-4155.
19. Salles G, Schuster SJ, Dreyling M, et al. Efficacy comparison of tisagenlecleucel vs usual care in patients with relapsed or refractory follicular lymphoma. Blood Adv. 2022;6(22):5835-5843.
20. Ghione P, Palomba ML, Patel A, et al. Comparative effectiveness of ZUMA-5 (axi-cel) vs SCHOLAR-5 external control in relapsed/refractory follicular lymphoma. Blood. 2022;140(8):851-860.