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

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

Received: June 14, 2023.
Accepted: November 17, 2023.


Publisher’s Disclaimer.
E-publishing ahead of print is increasingly important for the rapid dissemination of science. Haematologica is, therefore, E-publishing PDF files of an early version of manuscripts that have completed a regular peer review and have been accepted for publication. E-publishing of this PDF file has been approved by the authors. After having E-published Ahead of Print, manuscripts will then undergo technical and English editing, typesetting, proof correction and be presented for the authors’ final approval; the final version of the manuscript will then appear in a regular issue of the journal. All legal disclaimers that apply to the journal also pertain to this production process.
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. Maurer1,2, Carla Casulo3, Melissa C. Larson1, Thomas M. Habermann2, Izidore S. Lossos4, Yucai Wang2, Loretta J. Nastoupil5, Christopher Strouse6, Dai Chihara5, Peter Martin7, Jonathon B. Cohen8, Brad S. Kahl9, W Richard Burack3, Jean L. Koffa, Yong Mun10, Anthony Masaquel10, Mei Wu10, Michael C. Wei10, Ashwin Shewade10, Jia Li10, James R. Cerhan1, Brian K. Link6, Christopher R. Flowers5

1Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN, USA, 55902, 2Division of Hematology, Mayo Clinic, Rochester, MN, USA, 55905, 3Department of Medicine, Wilmot Cancer Institute, University of Rochester, Rochester, NY, USA, 14642, 4Department of Medicine, Comprehensive Sylvester Cancer Center, University of Miami, Miami, FL, USA, 33124, 5University of Texas, MD Anderson Cancer Center, Houston, TX, USA, 77030, 6Department of Medicine, University of Iowa, Iowa City, IA, USA, 52242, 7Department of Medicine, Weill Cornell Medical College, New York, NY, USA, 10075, 8Department of Hematology and Medical Oncology, Winship Cancer Institute, Emory University, Atlanta, GA, USA, 30322, 9Department of Medicine, Washington University in St Louis, St Louis, MO, USA, 63130, 10Genentech, Inc., South San Francisco, CA, USA, 94080

Correspondence to:
Dr. Matthew J. Maurer
Department of Quantitative Health Sciences
Mayo Clinic
200 First St SW, Rochester, MN 55905, USA.
Tel: +1-507-266-2064
E-mail: maurer.matthew@mayo.edu
ORCID ID: https://orcid.org/0000-0002-1867-0526

Running head: MAIC in 3L+ FL

Scientific category: Lymphoid Neoplasia

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.

Authorship Contributions:

• Conception and/or design of the study/research: MM, CC, ML, JL, TH, IL, YW, LN, CS, DC, PM, JC, BK, WB, JK, YM, AM, MW, MCW, AS, JL, JC, BL, CF

• Collection and/or assembly of the data: MM, CC, ML, JL, TH, IL, YW, LN, CS, DC, PM, JC, BK, WB, JK, JC, BL, CF
• Analysis and/or interpretation of the data: MM, CC, ML, JL, TH, IL, YW, LN, CS, DC, PM, JC, BK, WB, JK, YM, AM, MW, MCW, AS, JL, JC, BL, CF

• Wrote the manuscript: MM, CC, ML, JL, TH, IL, YW, LN, CS, DC, PM, JC, BK, WB, JK, YM, AM, MW, MCW, AS, JL, JC, BL, CF

Disclosures:

MJM: advisory board (Adaptive Biotechnologies, AstraZeneca, GenMab); consulting (BMS); research funding (Genmab, Genentech/Roche, BMS); spouse employment and stock (Exact Sciences).

CC: research funding (SecuraBio, Genmab, Gilead, Genentech, Inc.).

MCL: research funding (Genentech, Genmab).

TMH: 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: current employment (University of Miami), honoraria (Adaptive); consulting or advisory role (LRF); research funding (NCI); provided compensation teaching (Kyowa Kirin).

YW: 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: 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: current employment (University of Iowa); honoraria (Iowa Oncology Society, Curio Science); consulting or advisory role (Pfizer, GSK).

DC: research funding (Janssen, Epizyme, BMS, MorphoSys)

PM: consulting or advisory role (Abbvie, AztraZeneca, Beigene, BMS, Daiichi, Genentech, Janssen).

JBC: 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 Therapeutics, BiolInvent); Expert testimony (Hawaii).

BSK: consultancy (Astrazeneca, ADCT Therapeutics, Roche, Genetech, Abbvie, MEI, AcertaPharma,Pharmaceuticals, 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).

WRB: has no conflicts to declare.

JLK: consulting or advisory role (Gamida Cell, MorphoSys, BeiGene USA, TG Therapeutics); research funding (Viracta Therapeutics).

YM: current employment (Genentech/Roche); stock ownership (Genentech/Roche).

AM: current holder of individual stocks in a privately-held company (Roche); Current holder of stock options in a privately-held company (Roche).

MCW: current employment (Genentech); stock ownership (Roche); travel, accommodation, and expenses (Genentech/Roche).

AS: current employment (F. Hoffmann-La Roche Ltd); stock ownership (F. Hoffmann-La Roche Ltd).

JL: stock ownership (F. Hoffmann-La Roche Ltd).

JC: consulting or advisory role (BMS, Protagonist Therapeutics SMC).

BKL: consulting or advisory role (Genentech/Roche; MEI Pharma Inc); research funding (Genentech/Roche; Genmab; Celgene; AstraZeneca); expert testimony (Amgen).

CRF: 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).

Data sharing

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/
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 of 80%, complete response rate of 60%, and a median progression-free survival 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. Overall response rates (73%, 95% CI:65-80%) and complete response 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). Progression-free survival at 12 months was similar in the weighted LEO CReWE (60%, 95% CI:51-69%) and the mosunetuzumab trial (PFS 58%, 95% CI:47-68%). 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 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 clinical unmet 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 IRB. 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 (list in supplemental material) 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.

Matching variables were as follows: Age (years) at index therapy (mean, SD); prior lines of
therapy (mean, SD); progression of disease within 24 months (POD24) following frontline IC
(yes vs no vs did not receive IC); double-refractory to anti-CD20 and alkylator therapy (yes vs
no); and elevated LDH at index therapy (yes vs no). Matching-adjusted indirect comparison\(^\text{10}\)
(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 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 supplemental methods.

**Statistical methods**

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). Matching-adjusted indirect comparison\(^\text{10}\), a form of propensity score
weighting, was applied to individual patient data (IPD) from the LEO CReWE study. A series of
sensitivity analyses were 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, 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. Associations between groups and
categorical endpoints (e.g., ORR and CR) were assessed using logistic regression and
summarized with odds ratios and 95% confidence intervals. Further details on the study
analyses can be found in the SAP (supplemental material).

**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 (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 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, sixty-three patients (31%) had disease progression after 24 months to frontline IC, and 53 patients (26%) did not receive first line IC. One hundred twenty-three patients (61%) had received prior anthracycline, and 26 patients (13%) had received a prior autologous stem cell transplant. One hundred 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%) (Table S3.) Twenty-eight patients (14%) received a stem cell transplant as index therapy (N=18 autologous and N=10 allogenic), 7 received CAR-T, and 4 received a bi-specific antibody; 80 patients (40%) received index therapy on a clinical trial (Table S4). Most patients (N=171, 85%) received their index therapy at one of the 8 LEO institutions. Response assessment on index therapy was PET- or PET/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 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 reported8. 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 frontline immunochemotherapy, and 48 (53%) had double-refractory disease to previous anti-CD20 and alkylator therapies.

Primary comparison of LEO CReWE to GO29781 (MAIC 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 Supplemental Figure 1. 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%: 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 effective (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 small effects 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, Supplemental Figure 2), 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, Supplemental Figure 3. 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%), 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 centers 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\textsuperscript{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 to 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 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 in patients who had prior bone marrow involvement, and clinical plans to proceed to transplantation based on response may influence decision making. In the GO29781 trial, a bone marrow biopsy (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 NCTN clinical trials of follicular lymphoma, Rutherford and colleagues identified that BMB were irrelevant to assessing complete response in 99% of patients\textsuperscript{15}. Thus it is unlikely that differences in routine BMB between trial and observational cohorts has a significant impact on interpretation of results. Unlike aggressive lymphoma, where progression often manifests clinically prior to planned imaging\textsuperscript{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 within LEO CReWE identified that 54-64% of relapses to frontline therapy for FL were detected clinically\textsuperscript{17}. A distinct advantage of the LEO consortium is the involvement of experts in lymphoma clinical trials and patient management, including leaders in the national clinical trials network (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\textsuperscript{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\textsuperscript{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 with relapsed/refractory 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 pretreated 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.
REFERENCES


# Tables

**Table 1: Comparison of Key Patient Characteristics in GO29781 vs LEO CReWE Cohort (Unweighted and MAIC Weighted)**

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

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

<table>
<thead>
<tr>
<th>Group</th>
<th>N (Evaluable for Response)</th>
<th>ORR (95% CI)</th>
<th>CR Rate (95% CI)</th>
<th>PFS12 (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEO CReWE (unweighted)</td>
<td>202 (192)</td>
<td>77.6 (70.9-83.2)</td>
<td>57.8 (50.5-64.8)</td>
<td>65.0 (58.6-72.2)</td>
</tr>
<tr>
<td>LEO CReWE (MAIC Weighted)</td>
<td>167 (160)</td>
<td>73.0 (65.3-79.5)</td>
<td>52.9 (44.8-60.7)</td>
<td>59.5 (51.0-69.3)</td>
</tr>
<tr>
<td>GO29781 (trial results)</td>
<td>90 (90)</td>
<td>80.0 (70.3-87.7)</td>
<td>60.0 (49.1-70.2)</td>
<td>57.7 (46.9-68.4)</td>
</tr>
</tbody>
</table>

ORR=overall response rate; CR=complete response; PFS12=progression free survival at 12 months
Table 3: Sensitivity Analyses of MAIC Weighting Scenarios in LEO CReWE Cohort

<table>
<thead>
<tr>
<th>Scenario</th>
<th>N</th>
<th>ESS</th>
<th>MAIC Weighted ORR (95% CI)</th>
<th>MAIC Weighted CR (95% CI)</th>
<th>MAIC Weighted PFS12 (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unweighted Analysis</td>
<td>202</td>
<td></td>
<td>77.6 (70.9,83.2)</td>
<td>57.8 (50.5,64.8)</td>
<td>65.0 (58.6,72.2)</td>
</tr>
<tr>
<td>Primary MAIC Analysis</td>
<td>202</td>
<td>127</td>
<td>73.0 (65.3,79.5)</td>
<td>52.9 (44.8,60.7)</td>
<td>59.5 (51.0,69.3)</td>
</tr>
</tbody>
</table>

Change Trial I/E criteria application

<table>
<thead>
<tr>
<th>Change</th>
<th>No trial I/E criteria</th>
<th>357</th>
<th>297</th>
<th>73.7 (68.2,78.6)</th>
<th>56.0 (50.1,61.8)</th>
<th>63.1 (57.3,69.5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 LEO clinician trial I/E criteria</td>
<td>217</td>
<td>172</td>
<td>74.1 (67.0,80.2)</td>
<td>54.6 (47.1,61.9)</td>
<td>64.7 (57.3,73.0)</td>
<td></td>
</tr>
</tbody>
</table>

Change Matching variables

<table>
<thead>
<tr>
<th>Change</th>
<th>N</th>
<th>ESS</th>
<th>MAIC Weighted ORR (95% CI)</th>
<th>MAIC Weighted CR (95% CI)</th>
<th>MAIC Weighted PFS12 (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 Add gender</td>
<td>202</td>
<td>128</td>
<td>72.9 (65.2,79.5)</td>
<td>52.3 (44.3,60.2)</td>
<td>59.2 (50.8,69.1)</td>
</tr>
<tr>
<td>4 Add stage</td>
<td>190</td>
<td>116</td>
<td>70.6 (62.4,77.7)</td>
<td>52.0 (43.6,60.3)</td>
<td>59.5 (50.7,69.8)</td>
</tr>
<tr>
<td>5 Add ECOG PS</td>
<td>183</td>
<td>114</td>
<td>73.6 (65.5,80.4)</td>
<td>52.7 (44.2,61.0)</td>
<td>62.1 (53.2,72.5)</td>
</tr>
<tr>
<td>6 Add FLIPI</td>
<td>188</td>
<td>117</td>
<td>69.3 (61.0,76.5)</td>
<td>49.1 (40.8,57.4)</td>
<td>56.5 (47.8,66.8)</td>
</tr>
<tr>
<td>7 Add Bulky disease</td>
<td>191</td>
<td>120</td>
<td>70.4 (62.4,77.4)</td>
<td>52.8 (44.6,60.8)</td>
<td>55.6 (47.0,65.9)</td>
</tr>
<tr>
<td>8 Substitute refractory to prior line</td>
<td>203</td>
<td>132</td>
<td>71.7 (64.1,78.4)</td>
<td>51.1 (43.2,59.0)</td>
<td>58.4 (50.1,68.1)</td>
</tr>
<tr>
<td>9 Substitute POD24 to any 1L</td>
<td>201</td>
<td>127</td>
<td>74.4 (66.9,80.8)</td>
<td>53.3 (45.3,61.1)</td>
<td>59.9 (51.4,69.7)</td>
</tr>
<tr>
<td>10 Dichotomize prior LOT</td>
<td>202</td>
<td>133</td>
<td>66.7 (58.7,73.9)</td>
<td>50.6 (42.6,58.6)</td>
<td>55.3 (47.0,64.9)</td>
</tr>
</tbody>
</table>

Change Selection of Index Therapy

<table>
<thead>
<tr>
<th>Change</th>
<th>N</th>
<th>ESS</th>
<th>MAIC Weighted ORR (95% CI)</th>
<th>MAIC Weighted CR (95% CI)</th>
<th>MAIC Weighted PFS12 (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>11 Random line</td>
<td>202</td>
<td>85</td>
<td>71.5 (63.1,78.8)</td>
<td>41.7 (33.4,50.5)</td>
<td>54.1 (43.8,66.8)</td>
</tr>
<tr>
<td>12 First eligible line</td>
<td>202</td>
<td>20</td>
<td>49.4 (36.5,62.5)</td>
<td>34.8 (23.0,48.5)</td>
<td>34.4 (21.2,55.8)</td>
</tr>
<tr>
<td>13 Last eligible line</td>
<td>202</td>
<td>126</td>
<td>73.7 (66.1,80.2)</td>
<td>53.1 (45.1,61.0)</td>
<td>61.5 (53.1,71.3)</td>
</tr>
</tbody>
</table>

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; ORR=overall response rate; CR=complete response; PFS12=progression free survival at 12 months; I/E=inclusion/exclusion; ECOG PS=ECOG performance status
Figure 1: Analysis Flow Chart

Figure 2: Kaplan-Meier Curve of Progression-Free Survival for LEO CReWE vs GO29781 Trial
Patients receiving 3L+ with prior Alkylator and CD20 therapy from LEO CReWE [cite], N=441 Patients

Application of I/E Criteria from GO29781 trial, N=211 patients

Restrict to those with non-missing matching variables, N=202 patients

Select a single index line for each patient in LEO CReWE cohort

Final LEO CReWE dataset for analysis

Summarize outcomes in unweighted LEO CReWE dataset

Generate patient case weights for LEO CReWE patients via MAIC to rebalance LEO CReWE cohort to better reflect patient characteristics of GO29781

Summarize outcomes in weighted LEO CReWE dataset and compare to GO29781
Supplemental methods

Retrospective data capture allowed information to be captured on all lines of therapy prior to the cut-off date. Thus, a given patient may have had multiple lines of therapy that qualified as eligible for inclusion in the primary analysis. To identify the line of therapy utilized (i.e., index line) for each patient that most closely aligned with the clinical characteristics of the patients on the GO29781 study, preliminary matching-adjusted indirect comparison (MAIC) weighting was performed on the LEO CReWE dataset, which included all potential lines of therapy for a patient. For any LEO patient with multiple eligible lines, the eligible line of therapy with the highest preliminary MAIC weight was selected as the index line for inclusion in the primary analysis. The MAIC weighting was then re-run on the LEO CReWE dataset containing the (single) selected index line for each patient to generate the final MAIC weights for comparison to the GO29781 study.

Matching variables were selected based on input regarding clinical relevance and importance from LEO clinicians and specified prior to analyses. Matching variables were as follows: Age (years) at index therapy (mean, SD); prior lines of therapy (mean, SD); progression of disease within 24 months (POD24) following frontline IC (yes vs no vs did not receive IC); double-refractory to anti-CD20 and alkylator therapy (yes vs no); and elevated LDH at index therapy (yes vs no). Refractory status was defined as failure to achieve CR or partial response (PR) following a given therapy or progression within 6 months after the end of the given therapy.

Study endpoints

CR rate was defined as the proportion of patients with a best documented response of CR. Bone marrow biopsy was performed as clinically indicated by the treating physician and was not required for response assessment in the LEO CReWE Cohort. PFS was defined as time from index date until progression or death from any cause. Patients without progression or death were censored at the date of last disease assessment or initiation of subsequent therapy, whichever was earlier. Due to limited follow-up on the GO29781 trial, overall survival was not evaluated in this analysis.

Statistical methods

Matching-adjusted indirect comparison\(^{10}\), a form of propensity score weighting, was applied to individual patient data (IPD) from the LEO CReWE study. This approach to estimating the propensity score was utilized to address our pre-specified principles of CReWE IPD being analyzed by LEO investigators and statisticians, GO29781 IPD being analyzed by industry collaborators, and only aggregate data being shared between the groups. Patients in the LEO CReWE were weighted by their inverse odds of being in the LEO CReWE dataset vs. the GO29781 dataset (e.g., propensity score) to balance the covariate distribution on the GO29781 trial. The MAIC weights were generated using the R package maic (https://cran.r-project.org/web/packages/maic/index.html), which is based on the process demonstrated in the UK National Institute for Health and Care Excellence Decision Support Unit \(^{11}\) Effective sample size (ESS) was calculated as previously described. Case-weighted versions of statistical
techniques (e.g., weighted Kaplan Meier, logistic regression, Cox models) were then applied to the LEO CReWE data using the weights from the MAIC.

Summary tables and figures from the GO29781 study were from the study publication; independent review committee assessed response rates are utilized in this analysis. Genentech/Roche provided high-resolution versions of figures to extract Kaplan-Meier estimates of survival curves. Genentech/Roche did not have access to patient level data from the LEO CReWE study.

All analyses were performed using R v4.0.3 unless otherwise specified.
Table S1: Inclusion/exclusion criteria applied to the LEO CReWE Cohort based on the GO29781 clinical trial criteria

A) Inclusion

a. Age $\geq$ 18 years  
b. Grades 1-3a FL  
c. Relapsed after or failed to respond to at least two prior lines of systemic therapy and must have received prior treatment with an anti-CD20-directed therapy and an alkylating agent.  
d. AST and ALT $\leq$ 3 x the ULN  
e. Total bilirubin $\leq$ 1.5 x ULN  
f. Platelet count $\geq$ 75,000/mm$^3$  
g. ANC $\geq$ 1000/mm$^3$  
h. Total hemoglobin $\geq$ 10 g/dL  
i. Serum creatinine $\leq$ ULN or estimated creatinine CL $\geq$ 60 mL/min by Cockcroft-Gault method  
j. At least one bi-dimensionally measurable lesion (> 1.5 cm in its largest dimension for nodal lesions, or > 1.0 cm in its largest dimension for extranodal lesions by computerized tomography [CT] scan or MRI)

B) Exclusion

a. Prior treatment with systemic immunotherapeutic agents for which the mechanism of action involves T cells, including but not limited to cytokine therapy and anti-CTLA-4, anti-PD-1 and anti-PD-L1 therapeutic antibodies, within 12 weeks  
b. Significant treatment-emergent immune-related adverse events  
c. Treatment with radiotherapy within 2 weeks prior  
d. Autologous SCT within 100 days prior  
e. Prior treatment with CAR-T therapy within 30 days before  
f. Prior allogeneic SCT  
g. Current or past history of CNS lymphoma  
h. Significant comorbidity/medical history (including solid organ transplant) that may confound treatment outcomes  
i. Received systemic immunosuppressive medications (including but not limited to cyclophosphamide, azathioprine, methotrexate, thalidomide, and anti-tumor necrosis factor agents) with the exception of corticosteroid treatment $\leq$ 10 mg/day prednisone or equivalent within 2 weeks prior
<table>
<thead>
<tr>
<th>Variable</th>
<th>LEO CReWE (unweighted)</th>
<th>LEO CReWE (MAIC weighted)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=202</td>
<td>Weighted N=167</td>
</tr>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Age, years</td>
<td>60.1 (10.8)</td>
<td>60.3 (10.5)</td>
</tr>
<tr>
<td>Time since diagnosis, months</td>
<td>70.5 (38.9)</td>
<td>72.6 (40.2)</td>
</tr>
<tr>
<td>Time since most recent therapy (months)</td>
<td>18.6 (21.1)</td>
<td>14.8 (19.3)</td>
</tr>
<tr>
<td></td>
<td>N (%)</td>
<td>N (%)</td>
</tr>
<tr>
<td>Clinical Characteristics at Index therapy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age &gt;60 years</td>
<td>100 (50%)</td>
<td>87 (52%)</td>
</tr>
<tr>
<td>Sex, male</td>
<td>118 (58%)</td>
<td>97 (58%)</td>
</tr>
<tr>
<td>Bulky disease (&gt;7cm)</td>
<td>31 (16%)</td>
<td>20 (13%)</td>
</tr>
<tr>
<td>Stage III/IV</td>
<td>157 (84%)</td>
<td>134 (86%)</td>
</tr>
<tr>
<td>HGB ≥12</td>
<td>202 (100%)</td>
<td>167 (100%)</td>
</tr>
<tr>
<td>ECOG PS 2-4</td>
<td>8 (4%)</td>
<td>8 (5%)</td>
</tr>
<tr>
<td>Elevated LDH</td>
<td>59 (29%)</td>
<td>65 (39%)</td>
</tr>
<tr>
<td>Grade</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-2</td>
<td>154 (76%)</td>
<td>128 (76%)</td>
</tr>
<tr>
<td>3A</td>
<td>34 (17%)</td>
<td>29 (17%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>14 (7%)</td>
<td>10 (6%)</td>
</tr>
<tr>
<td>FLIPI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>49 (24%)</td>
<td>32 (19%)</td>
</tr>
<tr>
<td>2</td>
<td>74 (37%)</td>
<td>54 (33%)</td>
</tr>
<tr>
<td>3-5</td>
<td>61 (30%)</td>
<td>67 (40%)</td>
</tr>
<tr>
<td>Treatment History Prior to Index Line</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>----------------</td>
</tr>
<tr>
<td>POD24 to 1L IC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>86 (43%)</td>
<td>70 (42%)</td>
</tr>
<tr>
<td>No</td>
<td>63 (31%)</td>
<td>70 (42%)</td>
</tr>
<tr>
<td>Other (non-IC) 1L therapy</td>
<td>53 (26%)</td>
<td>27 (16%)</td>
</tr>
<tr>
<td>Anti-CD20 refractory</td>
<td>141 (70%)</td>
<td>131 (78%)</td>
</tr>
<tr>
<td>Alkylator refractory</td>
<td>76 (38%)</td>
<td>90 (54%)</td>
</tr>
<tr>
<td>Double refractory</td>
<td>73 (36%)</td>
<td>89 (53%)</td>
</tr>
<tr>
<td>Prior LOT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>116 (57%)</td>
<td>82 (49%)</td>
</tr>
<tr>
<td>3</td>
<td>48 (24%)</td>
<td>33 (20%)</td>
</tr>
<tr>
<td>&gt;3</td>
<td>38 (19%)</td>
<td>53 (31%)</td>
</tr>
<tr>
<td>Refractory to most recent therapy</td>
<td>140 (69%)</td>
<td>129 (77%)</td>
</tr>
<tr>
<td>Prior SCT</td>
<td>26 (13%)</td>
<td>25 (15%)</td>
</tr>
</tbody>
</table>

<p>| Treatment Details on Index Line               |  |  |
| Class of therapy at index line               |  |  |
| CD20 Immunochemotherapy based                | 57 (28%)        | 36 (21%)       |
| CD20 lenalidomide based                      | 28 (14%)        | 24 (15%)       |
| Platinum based salvage                       | 23 (11%)        | 18 (11%)       |
| CD20 monotherapy                             | 18 (9%)         | 11 (7%)        |
| Radioimmunotherapy based                     | 9 (4%)          | 6 (4%)         |
| CAR-T                                         | 7 (3%)          | 6 (4%)         |
| Novel agent + CD20 antibody                  | 16 (8%)         | 11 (7%)        |
| Novel agent monotherapy                      | 31 (15%)        | 42 (25%)       |</p>
<table>
<thead>
<tr>
<th></th>
<th>Value 1</th>
<th>Value 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other</td>
<td>13 (6%)</td>
<td>13 (8%)</td>
</tr>
<tr>
<td>Treatment location at LEO center</td>
<td>171 (85%)</td>
<td>142 (85%)</td>
</tr>
<tr>
<td>Treated on clinical trial at index therapy</td>
<td>80 (40%)</td>
<td>76 (45%)</td>
</tr>
<tr>
<td><strong>Response assessment modality</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PET or PET/CT</td>
<td>109 (54%)</td>
<td>88 (52%)</td>
</tr>
<tr>
<td>CT</td>
<td>73 (36%)</td>
<td>62 (37%)</td>
</tr>
<tr>
<td>Other</td>
<td>7 (3%)</td>
<td>6 (4%)</td>
</tr>
<tr>
<td>Unknown/missing</td>
<td>13 (6%)</td>
<td>11 (7%)</td>
</tr>
</tbody>
</table>
Figure S1: Distribution of Weights in Primary Analysis
**Post-hoc Analyses**

The following analyses were generated after the locked SAP in response to reviewer comments as part of the manuscript peer review process.

1) **Table S3.** Response Rates by Treatment Group Subsets in Primary Analysis

<table>
<thead>
<tr>
<th>Treatment Group</th>
<th>Total (N=202)</th>
<th>Unweighted %</th>
<th>MAIC weighted %</th>
<th>Unweighted ORR (95% CI)</th>
<th>Unweighted CR Rate (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD20 immunochemotherapy based</td>
<td>57</td>
<td>28.2%</td>
<td>21.3%</td>
<td>94.5 (83.9, 98.6)</td>
<td>78.2 (64.6, 87.8)</td>
</tr>
<tr>
<td>CD20 lenalidomide based</td>
<td>28</td>
<td>13.9%</td>
<td>14.6%</td>
<td>66.7 (44.7, 83.6)</td>
<td>29.2 (13.4, 51.2)</td>
</tr>
<tr>
<td>Platinum based salvage</td>
<td>23</td>
<td>11.4%</td>
<td>10.7%</td>
<td>82.6 (60.5, 94.3)</td>
<td>60.9 (38.8, 79.5)</td>
</tr>
<tr>
<td>CD20 monotherapy</td>
<td>18</td>
<td>8.9%</td>
<td>6.5%</td>
<td>83.3 (57.7, 95.6)</td>
<td>72.2 (49.4, 89.3)</td>
</tr>
<tr>
<td>Radioimmunotherapy based</td>
<td>9</td>
<td>4.5%</td>
<td>3.8%</td>
<td>77.8 (40.2, 96.1)</td>
<td>55.6 (22.7, 84.7)</td>
</tr>
<tr>
<td>CAR-T</td>
<td>7</td>
<td>3.5%</td>
<td>3.7%</td>
<td>85.7 (42.0, 99.2)</td>
<td>85.7 (42.0, 99.2)</td>
</tr>
<tr>
<td>Novel agent + CD20</td>
<td>16</td>
<td>7.9%</td>
<td>6.7%</td>
<td>84.6 (53.7, 97.3)</td>
<td>76.9 (46.0, 93.8)</td>
</tr>
<tr>
<td>Novel agent monotherapy</td>
<td>31</td>
<td>15.3%</td>
<td>25.2%</td>
<td>40.0 (23.2, 59.2)</td>
<td>23.3 (10.6, 42.7)</td>
</tr>
<tr>
<td>Other</td>
<td>13</td>
<td>6.4%</td>
<td>7.5%</td>
<td>84.6 (53.7, 97.3)</td>
<td>46.2 (20.4, 73.9)</td>
</tr>
</tbody>
</table>
2) **Table S4.** Treatment Group Subsets by Clinical Trial Status in Primary Analysis

<table>
<thead>
<tr>
<th>Index Therapy Treatment Group</th>
<th>On Clinical Trial (N=80)</th>
<th>Not on Clinical Trial (N=122)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>CD20 immunochemotherapy based</td>
<td>10</td>
<td>12.5%</td>
</tr>
<tr>
<td>CD20 lenalidomide based</td>
<td>7</td>
<td>8.8%</td>
</tr>
<tr>
<td>Platinum based salvage</td>
<td>3</td>
<td>3.8%</td>
</tr>
<tr>
<td>CD20 monotherapy</td>
<td>1</td>
<td>1.2%</td>
</tr>
<tr>
<td>Radioimmunotherapy based</td>
<td>3</td>
<td>3.8%</td>
</tr>
<tr>
<td>CAR-T</td>
<td>7</td>
<td>8.8%</td>
</tr>
<tr>
<td>Novel agent + CD20</td>
<td>16</td>
<td>20.0%</td>
</tr>
<tr>
<td>Checkpoint inhibitor</td>
<td>4</td>
<td>5.0%</td>
</tr>
<tr>
<td>IMID</td>
<td>3</td>
<td>3.8%</td>
</tr>
<tr>
<td>PI3K</td>
<td>3</td>
<td>3.8%</td>
</tr>
<tr>
<td>Other N-of-1</td>
<td>6</td>
<td>7.5%</td>
</tr>
<tr>
<td>Novel agent monotherapy</td>
<td>26</td>
<td>32.5%</td>
</tr>
<tr>
<td>PI3K</td>
<td>5</td>
<td>6.3%</td>
</tr>
<tr>
<td>Bispecific antibody</td>
<td>4</td>
<td>5.0%</td>
</tr>
<tr>
<td>Checkpoint inhibitor</td>
<td>3</td>
<td>3.8%</td>
</tr>
<tr>
<td>BTKi</td>
<td>3</td>
<td>3.8%</td>
</tr>
<tr>
<td>CD19</td>
<td>2</td>
<td>2.5%</td>
</tr>
<tr>
<td>CD22</td>
<td>2</td>
<td>2.5%</td>
</tr>
<tr>
<td>Other N-of-1</td>
<td>7</td>
<td>8.8%</td>
</tr>
<tr>
<td>Other</td>
<td>7</td>
<td>8.8%</td>
</tr>
<tr>
<td>Chemotherapy without CD20</td>
<td>1</td>
<td>1.3%</td>
</tr>
<tr>
<td>Checkpoint inhibitor based combinations</td>
<td>3</td>
<td>3.8%</td>
</tr>
<tr>
<td>Other N-of-1</td>
<td>3</td>
<td>3.8%</td>
</tr>
</tbody>
</table>
3) **Table S5.** Outcomes in Primary Analysis by Receipt of Index Therapy on a Clinical Trial

<table>
<thead>
<tr>
<th>Index therapy on clinical trial</th>
<th>N</th>
<th>MAIC Weighted ORR (95% CI)</th>
<th>MAIC Weighted CR (95% CI)</th>
<th>MAIC Weighted PFS12 (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>80</td>
<td>76.3 (64.7, 85.1)</td>
<td>56.0 (44.0, 67.4)</td>
<td>56.9 (44.4, 72.9)</td>
</tr>
<tr>
<td>No</td>
<td>122</td>
<td>70.2 (59.2, 79.3)</td>
<td>50.1 (39.7, 60.6)</td>
<td>61.9 (51.0, 75.1)</td>
</tr>
</tbody>
</table>

4) **TTNT**

Time to next treatment was defined as the time from index therapy until the initiation of the subsequent line of treatment. The MAIC weighted time to next treatment at 12 months (TTNT12) was 69.5%, 95% CI: (60.9-79.4). The TTNT at 12 months on the GO29781 trial was 68.1%, 95% CI: (58.3-77.9).

5) **Figure S2.** Distribution of weights for random line matching (Table 3, scenario 11). A much higher percentage of near-zero weights was observed in scenario 11 (ESS=84.7) than the primary analysis.
6) **Figure S3.** Distribution of weights for first eligible line matching (Table 3, scenario 12). Nearly all index therapies received near-zero weights with a small number of highly influential cases contributing most of the information (ESS=20.2).
TITLE: Comparison of Outcomes in Clinical Trial GO29781 to LEO CREWE-FL Patients with Relapsed/Refractory Follicular Lymphoma Receiving Three or More Lines of Systemic Therapy

PROJECT NUMBER: LYMLEO2001

VERSION NUMBER: 1.0

PLAN PREPARED BY: Matthew J. Maurer, MS DMSc, Melissa C. Larson, MS
LEO Cohort Statistics and Informatics Core

DATE FINAL: {2 February 2022}

RESEARCH INITIATOR
Matthew J. Maurer, MS, DMSc
Division of Clinical Trials and Biostatistics
Mayo Clinic, Rochester, MN

STATISTICAL ANALYSIS PLAN For LEO Cohort Projects
1. Background and Objectives

1.1 Background

Non-Hodgkin Lymphoma (NHL) is one of the leading causes of cancer death in the US\(^1\) and in Europe\(^2\). Follicular Lymphoma (FL) is the second most common type of lymphoma diagnosed in the US and Western Europe accounting for approximately 20-30% of all NHL cases.\(^3\) Approximately 80-85% of all NHL cases have B-cell origin. Among these, FL has a long natural history and median OS of over 10 years.\(^4\) Treatment regimens consisting of B-cell targeting monoclonal antibodies combined with chemotherapy have been the mainstay of treatment and have led to improvements in response rates and survival.\(^5,6\) While FL is an indolent disease, it is associated with frequent relapses. Additionally, a significantly higher risk of death is associated with early progression of the disease.\(^7,8\) Progression of disease after frontline treatment typically results in shorter disease-free intervals and increased refractoriness with each subsequent progression/relapse.\(^9,10\) Patients who have received two or more lines of prior systemic therapy have poor prognosis with a median progression-free survival (PFS) of approximately 1 year to 3\(^{rd}\) line therapies\(^10,11\).

Improving treatment in the relapsed/refractory (r/r) setting is a major need in FL. To address this need, we initiated the Lymphoma Epidemiology of Outcomes (LEO) Consortium to build a real-world cohort in this setting by bringing together the LEO Cohort study and institutional databases at LEO centers. This cohort provides real world data (RWD) which can be utilized as benchmark and/or external RWD control arm to provide context to evidence generated from single arm clinical studies to provide actionable, real world evidence (RWE) on the efficacy of new treatments.

1.2 Study Objectives

a) Generate a comparison cohort of LEO CREWE patients with r/r FL who have received at least two prior lines of systemic therapy, including alkylator and rituximab base therapy, aligning with key eligibility criteria for the Genentech GO29781 study.

b) Describe the clinical characteristics and outcomes of LEO CREWE patients meeting eligibility in 1.2.a

c) Perform a matching adjusted indirect comparison (MAIC) analysis of the LEO CREWE comparison cohort to the Genentech GO29781 study.

2. Research Design

2.1 Study Design

This is a non-interventional/observational study based on secondary use of data collected in the LEO Cohort as well as LEO Consortium institutional databases. All patient management, including treatment choices and follow-up (including use of re-biopsy and scans), are per clinical practice of the treating physician, and these details will be abstracted from documentation in the medical record at the LEO Consortium centers as well as outside medical records as needed. While the LEO Cohort was prospectively recruited and followed, details on r/r disease (beyond fact of occurrence, which is prospectively documented) need to be abstracted in a retrospective manner from medical records. For the institutional databases, while a variety of case identification approaches are being used, all data with respect to r/r events will also be abstracted retrospectively. Thus, all cases for this effort will be
abstracted from medical records across the LEO centers to build a database to ensure consistency and comparability across data sources.

2.2 Data Sources

The Lymphoma Epidemiology of Outcomes (LEO) Consortium was used to build a RWD cohort in the setting for r/r FL with at least two prior lines of systemic therapy by bringing together the LEO Cohort Study and institutional databases at LEO centers. The LEO academic cancer centers include Cornell University, Emory University, Mayo Clinic Rochester, MD Anderson, University of Iowa, University of Miami, University of Rochester, and Washington University. Given the slow-progressing and indolent nature of FL, patients in the prospectively enrolled LEO Cohort will be supplemented with patients with r/r FL from the individual databases at each academic site within the LEO Consortium in order to achieve an adequate sample size for the proposed descriptive benchmark in 3L and higher FL. For the institutional databases, two major approaches were utilized: 1) cases who were first diagnosed at the LEO institution in 2002 or later and followed by the center (subsequent clinical course will be abstracted, including outside records as feasible); and 2) cases who first appeared at the LEO institution from 2010-2018 with r/r disease and already had at least two prior lines of therapy (all therapies will be abstracted, including outside records as feasible). Full details on the LEO CReWE cohort are provided in the initial publication12.

2.3 Primary Outcome Measure

The primary outcome measure for this study will be as follows:

Overall response rate (ORR) defined as the proportion of patients with best response as complete response (CR) or partial response (PR) during the available follow-up beginning from the index date to 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. Response assessments will be based on documented clinical information in the patient’s chart and abstracted by a clinical reviewer using a standard protocol. Response may be documented as one of the following categories: CR, PR, stable disease (SD), progressive disease (PD), indeterminate or missing. The response variable utilized from the GO29781 trial will be the independent review facility (IRF) assessed response rate as this was defined to be the primary endpoint for the trial.

2.4 Secondary Outcome Measures

Secondary outcome measures for this study will be as follows:

1) Complete response (CR) rate as defined as proportion of patients with best documented response of CR during the available follow-up beginning from the index date to 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.
2) Progression-free survival (PFS) defined as time from index date until progression or death from any cause. Patients without progression or death will be censored at date of last disease assessment or initiation of subsequent therapy, whichever is earlier.

3) Overall survival, defined as time from index date until death from any cause. Patients without death will be censored at date of last follow-up.

2.5 Exploratory Outcome Measures

Exploratory outcome measures for this study will be as follows:

1) Duration of response (DOR) defined as the time from date of first response (CR or PR) to earliest of PD or documented progression or death from any cause in the subset of patients with response (PR or CR). Patients without progression or death will be censored at date of last disease assessment or initiation of subsequent therapy, whichever is earlier. This endpoint is exploratory due to the heterogeneity in response assessment in routine clinical care as compared to standard response measurements in a clinical trial.

2.6 Determination of Sample Size

This is a descriptive analysis of retrospective data without formally controlling for the type I error rate and therefore no formal calculation of sample size is performed.

It is expected that approximately 1200 patients in the LEO Consortium may have had at least two prior lines of systemic therapy for r/r FL. We estimate that 500 patients will have received systemic therapy for follicular lymphoma after two or prior lines and have received prior treatment with an alkylator based chemotherapy and an anti-CD20 antibody therapy. This will comprise the starting cohort for analysis. We further estimate that approximately 300 patients would meet all eligibility criteria for Genentech GO29781 study described in Appendix 1. This will comprise the cohort used for the primary MAIC analysis. Table 1 shows a shell to describe the expected precision in the estimate of unweighted ORR as the primary endpoint for this study. With an observed ORR of 60%, a sample size of 300 patients would result in a 95% confidence interval of 54% - 66%, while an observed ORR of 80% and a sample of size of 75 would result in a 95% confidence interval of 69% - 88%.

Table 1. Unweighted confidence intervals for ORR in the LEO-CReWE RWD cohort

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>True ORR</th>
<th>Clopper-Pearson 95% CI Limits</th>
<th>True ORR</th>
<th>Clopper-Pearson 95% CI Limits</th>
<th>True ORR</th>
<th>Clopper-Pearson 95% CI Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>75</td>
<td>60%</td>
<td>48%, 71%</td>
<td>70%</td>
<td>58%, 80%</td>
<td>80%</td>
<td>69%, 88%</td>
</tr>
<tr>
<td>100</td>
<td>60%</td>
<td>50%, 70%</td>
<td>70%</td>
<td>60%, 79%</td>
<td>80%</td>
<td>71%, 87%</td>
</tr>
<tr>
<td>200</td>
<td>60%</td>
<td>53%, 69%</td>
<td>70%</td>
<td>63%, 76%</td>
<td>80%</td>
<td>74%, 85%</td>
</tr>
<tr>
<td>300</td>
<td>60%</td>
<td>54%, 66%</td>
<td>70%</td>
<td>65%, 75%</td>
<td>80%</td>
<td>75%, 84%</td>
</tr>
</tbody>
</table>
2.7 Statistical Methods

For descriptive summaries, continuous variables will be summarized using descriptive statistics (i.e., N, mean, median, standard deviation, IQR, and range). Categorical variables, including response rates, will be summarized as proportions and/or rates. Time to event variables will be summarized using Kaplan-Meier curves. 95% confidence intervals will be provided for point estimates.

Associations between groups and categorical endpoints (e.g. ORR and CR) will be assessed using logistic regression and summarized with odds ratios and 95% confidence intervals. Associations between groups and time to event endpoints (e.g. PFS and DOR) will be assessed using Cox proportional hazards models and summarized with hazard ratios and 95% confidence intervals.

Weighted versions of the above statistical models (e.g. weighted Kaplan Meier, logistic regression, Cox models) will be utilized in the MAIC analysis, with weights determined by the MAIC matching in 2.9. All analyses will be performed using R v4.0.3 unless otherwise specified.

2.8 Analysis Population and Eligibility Criteria

The starting population for the study (e.g. Cohort 1 in prior publication12) will comprise of all patients meeting the following criteria:

a) received systemic therapy for FL grade 1-3A after at least two prior lines of systemic therapy
   a. Any therapies after transformation to grade 3B FL, DLBCL or high-grade B-cell will not be included in determination of eligibility and/or analyses
b) received prior treatment with an anti-CD20-directed therapy and an alkylating agent.

The retrospective setting allows information on all lines of therapy recorded in the patient’s chart up until the latest line of therapy recorded prior to the cut-off date for when the patient’s chart will be reviewed to abstract information regarding the primary analysis. This may mean that a given patient may have multiple eligible lines of therapy to be qualified as potential index regimens as part of the broad cohort. The index line for each patient will be determined for each analysis after applying any additional eligibility criteria as defined in the primary and sensitivity analysis below.

2.9 Primary Analysis

2.9.0 Data cutoff

i) All GO29781 results will utilize the August 27, 2021 data cutoff to align with the final analysis presented at the 2021 American Society of Hematology Annual Meeting and primary publication, unless otherwise specified.

ii) All LEO CReWE-FL analysis will utilize the locked dataset from the initial CReWE-FL publication12, unless otherwise specified.
2.9.1 Matching-adjusted indirect comparison (MAIC)\textsuperscript{13}, a form of propensity score weighting, will be applied to individual patient data (IPD) from the LEO CReWE study with aggregate data on pre-specified set of prognostic factors from the Phase I/IB study pivotal 3L+ FL expansion cohort (GO29781). This approach to estimating the propensity score is utilized due to IPD only being available in the LEO data, while aggregate data is available for the GO29781 trial. Briefly, patients in the LEO CReWE dataset will be weighted so the weighted mean baseline characteristics match those from the GO29781 trial. Patients in the LEO-CReWE (IPD) will be weighted by their inverse odds of being in the LEO CReWE dataset vs. the GO29781 dataset (e.g. propensity score) to balance the covariate distribution. Patient weights in the LEO-CReWE cohort will be rescaled so that a weight > 1 indicates an individual has more weight in the reweighted analysis population than the original unweighted population. Similarly, a weight < 1 indicates an individual has less weight in the reweighted analysis population than the original unweighted population.

The MAIC will be implemented using the R package maic (https://cran.r-project.org/web/packages/maic/index.html) which is based on the process demonstrated in the UK National Institute for Health and Care Excellence Decision Support Unit\textsuperscript{14}.

2.9.2 Eligibility: All patients from Cohort 1 with any lines of therapy that meet all of the following criteria:

a) All eligibility criteria from the GO29781 trial listed in Appendix 1 are met
b) All matching variables are non-missing

2.9.3 Matching variables: The following variables will be utilized for matching in the MAIC analysis. The variable list was developed based on input from expert LEO clinicians. The type of variable (continuous or categorized) is specified for each variable as follows:

1) Age (years) at index therapy (mean, SD)
2) Prior lines of therapy (mean, SD)
3) POD24 to frontline immunochemotherapy (yes vs no vs did not receive IC)
4) Double refractory to CD20 and alkylator therapy (yes vs no)
5) Elevated LDH at index therapy (yes vs no)

2.9.4 Selection of index therapy: It is expected that a subset of patients will have multiple lines of therapy that meet eligibility criteria above. For each patient with multiple eligible lines, the eligible line of therapy with the highest MAIC weight be utilized as the index line in the primary analysis.

2.9.5 Summary of MAIC analysis: The following results will be reported:

1) Flow diagram detailing the application of exclusion criteria and generation of the analysis dataset
2) Effective sample size and histogram of weights for each patient
3) Kaplan Meier curves of unweighted time to event outcomes from the LEO IPD data. LEO IPD curves will be truncated based on the observed follow-up from the GO29781 trial.
   a. PFS
   b. DOR
   c. OS
4) Kaplan Meier curves comparing weighted time to event outcomes from the LEO IPD data to the observed outcomes in the GO29781 trial. Curves for the GO29781 trial will be extracted from using DigitizeIt. LEO IPD curves will be truncated based on the observed follow-up from the GO29781 trial.
   a. PFS
   b. DOR
   c. OS

5) Weighted estimates and 95% CI from IPD data for
   a. ORR
   b. CR
   c. PFS at 12 months from start of therapy (e.g. PFS12)
   d. PFS at 18 months from start of therapy (e.g. PFS18)
   e. OS at 12 months from start of therapy (e.g. OS12)
   f. OS at 18 months from start of therapy (e.g. OS18)

6) Comparison of LEO CReWE-FL weighted estimates (reference) to GO29781 trial summary data
   a. Logistic regression, odds ratios, 95% CI, p-value
      i. ORR
      ii. CR

2.10 Sensitivity Analysis

A series of sensitivity analyses will be performed to assess the potential impact of missing data, eligibility criteria, choice of matching variables, and selection of index therapy. The following elements will be evaluated.

1) Choice of exclusion criteria
   a. All inclusion/exclusion criteria in Appendix 1 are applied and met (primary)
   b. LEO clinician trial eligibility criteria, as defined in the initial LEO-CReWE-FL manuscript\(^1\), are applied and met (sensitivity)
   c. Starting population (e.g. Cohort 1 in section 2.8) with no additional inclusion/exclusion criteria applied (sensitivity)

2) Choice of matching variables
   a. Age, prior lines of therapy, POD24 to 1L immunochemotherapy, double refractory to CD20/alkylator, LDH at index (primary)
   b. Additional variables to evaluate may include (sensitivity)
      i. Gender
      ii. Refractory to most prior therapy
      iii. POD24 to any 1L therapy
      iv. FLIPI at index therapy
      v. Stage at index therapy
      vi. ECOG PS at index therapy
      vii. Bulky disease at index therapy
      viii. Prior SCT at index therapy

3) Method of variable inclusion for MAIC analysis
a. Prior lines of therapy (mean, SD) (primary)
b. Prior lines of therapy (2 vs. >2) (sensitivity)
4) Selection of index therapy for patients with multiple eligible lines
   a. Highest MAIC weight as index line (primary)
   b. Random line as index line (sensitivity)
   c. First eligible line as index line (sensitivity)
   d. Last eligible line as index line (sensitivity)

2.11 Subgroup Analysis

Outcomes for the primary analysis may be reported in subsets of patients who received the following treatment groups at index therapy.

1) Immunochemotherapy
2) CD20 monotherapy
3) PI3K +/- CD20 therapy
4) Lenalidomide +/- CD20 therapy
5) Novel agent +/- CD20 therapy
6) Salvage and/or cellular therapy

2.12 Handling of Missing Data

The primary analysis will utilize complete case data. The impact of missing data on inclusion/exclusion criteria and requirement of non-missing data for matching variables will be summarized and evaluated as part of the sensitivity analysis. If necessary, imputation may be performed as part of additional sensitivity analyses.

2.13 Changes to SAP

The proposed analytical plan was developed and finalized apriori to the LEO Statistical Core performing any of the planned primary analysis. We may encounter issues with the implementation of the MAIC analysis using the proposed eligibility, treatment selection, and matching variables that result in a low-quality analysis (e.g. many patients being allocated near-zero weights, a small number of participants with very high weights, or imbalances in key clinical characteristics). This may necessitate modifications to the planned analysis, such as selection of a different set of matching variables. Any changes to the primary analysis will be discussed and approved by the study team and documented as modifications in the SAP.

2.14 Data Access and Study Deliverables

Data Access: All analyses will be performed by the LEO Statistics and Informatics Core. Genentech will not have access to patient level data from the LEO CReWE-FL study. Genentech will provide all necessary summary tables and figures from the GO29781 study to perform the MAIC matching and analysis.
Study deliverables: The LEO Statistics and Informatics Core will generate a study report from the proposed analysis in the SAP. This report will be provided to Genentech at the completion of the study. The LEO study team will develop and submit an academic manuscript from the study report.

3. References


4. Appendices
4.1 Appendix 1: Eligibility criteria for primary analysis

A) Inclusion
a. Age >= 18 years
b. Grades 1-3a FL
c. Relapsed after or failed to respond to at least two prior lines of systemic therapy and must have received prior treatment with an anti-CD20-directed therapy and an alkylating agent.
d. AST and ALT <= 3 x the ULN
e. Total bilirubin <= 1.5 x ULN
f. Platelet count >= 75,000/mm3
g. ANC >= 1000/mm3
h. Total hemoglobin >= 10 g/dL
i. Serum creatinine <= ULN or estimated creatinine CL >= 60 mL/min by Cockcroft-Gault method
j. At least one bi-dimensionally measurable lesion (> 1.5 cm in its largest dimension for nodal lesions, or > 1.0 cm in its largest dimension for extranodal lesions by computerized tomography [CT] scan or MRI)

B) Exclusion
a. Prior treatment with systemic immunotherapeutic agents for which the mechanism of action involves T cells, including but not limited to cytokine therapy and anti-CTLA-4, anti-PD-1 and anti-PD-L1 therapeutic antibodies, within 12 weeks
b. Significant treatment-emergent immune-related adverse events
c. Treatment with radiotherapy within 2 weeks prior
d. Autologous SCT within 100 days prior
e. Prior treatment with CAR-T therapy within 30 days before
f. Prior allogeneic SCT
g. Current or past history of CNS lymphoma
h. Significant comorbidity/medical history (including solid organ transplant) that may confound treatment outcomes
i. Received systemic immunosuppressive medications (including but not limited to cyclophosphamide, azathioprine, methotrexate, thalidomide, and anti-tumor necrosis factor agents) with the exception of corticosteroid treatment <= 10 mg/day prednisone or equivalent within 2 weeks prior