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 G029781 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 >= 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

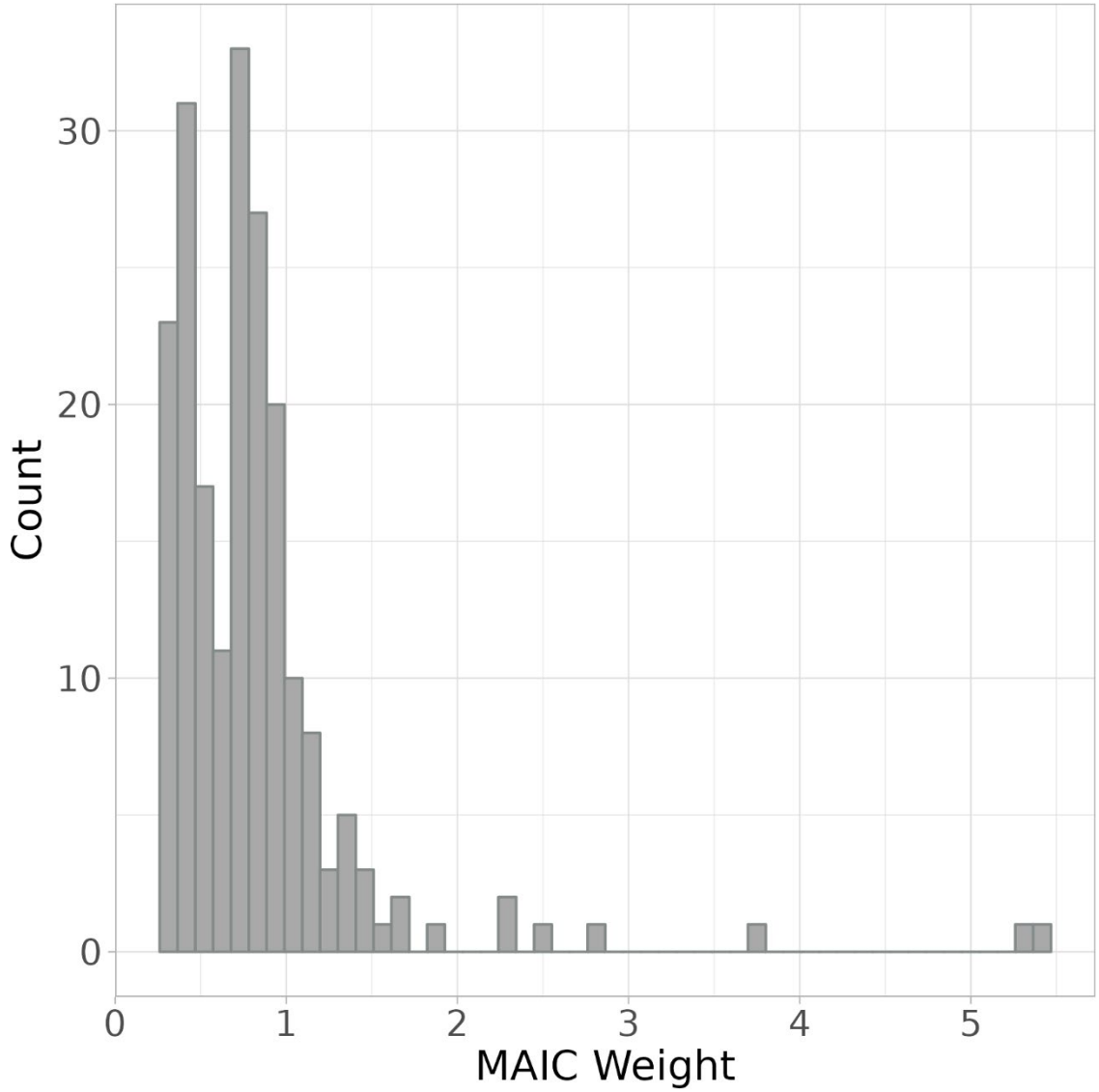
Table S2: Clinical details on the 202 LEO CReWE patients in primary analysis

Variable	LEO CReWE (unweighted) N=202	LEO CReWE (MAIC weighted) Weighted N=167
	Mean (SD)	Mean (SD)
Age, years	60.1 (10.8)	60.3 (10.5)
Time since diagnosis, months	70.5 (38.9)	72.6 (40.2)
Time since most recent therapy (months)	18.6 (21.1)	14.8 (19.3)
	N (%)	N (%)
Clinical Characteristics at Index therapy		
Age >60 years	100 (50%)	87 (52%)
Sex, male	118 (58%)	97 (58%)
Bulky disease (>7cm)	31 (16%)	20 (13%)
Stage III/IV	157 (84%)	134 (86%)
HGB ≥12	202 (100%)	167 (100%)
ECOG PS 2-4	8 (4%)	8 (5%)
Elevated LDH	59 (29%)	65 (39%)
Grade 1-2 3A Unknown	 154 (76%) 34 (17%) 14 (7%)	 128 (76%) 29 (17%) 10 (6%)
FLIPI 0-1 2 3-5	 49 (24%) 74 (37%) 61 (30%)	 32 (19%) 54 (33%) 67 (40%)

Unavailable	18 (9%)	14 (8%)
Treatment History Prior to Index Line		
POD24 to 1L IC		
Yes	86 (43%)	70 (42%)
No	63 (31%)	70 (42%)
Other (non-IC) 1L therapy	53 (26%)	27 (16%)
Anti-CD20 refractory	141 (70%)	131 (78%)
Alkylator refractory	76 (38%)	90 (54%)
Double refractory	73 (36%)	89 (53%)
Prior LOT		
2	116 (57%)	82 (49%)
3	48 (24%)	33 (20%)
>3	38 (19%)	53 (31%)
Refractory to most recent therapy	140 (69%)	129 (77%)
Prior SCT	26 (13%)	25 (15%)
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%)

Other	13 (6%)	13 (8%)
Treatment location at LEO center	171 (85%)	142 (85%)
Treated on clinical trial at index therapy	80 (40%)	76 (45%)
Response assessment modality		
PET or PET/CT	109 (54%)	88 (52%)
CT	73 (36%)	62 (37%)
Other	7 (3%)	6 (4%)
Unknown/missing	13 (6%)	11 (7%)

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

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

2) **Table S4.** Treatment Group Subsets by Clinical Trial Status in Primary Analysis

Index Therapy Treatment Group	On Clinical Trial (N=80)		Not on Clinical Trial (N=122)	
	N	%	N	%
CD20 immunochemotherapy based	10	12.5%	47	38.5%
CD20 lenalidomide based	7	8.8%	21	17.2%
Platinum based salvage	3	3.8%	20	16.4%
CD20 monotherapy	1	1.2%	17	13.9%
Radioimmunotherapy based	3	3.8%	6	4.9%
CAR-T	7	8.8%	0	0.0%
Novel agent + CD20	16	20.0%	0	0.0%
Checkpoint inhibitor	4		0	
IMID	3		0	
PI3K	3		0	
Other N-of-1	6		0	
Novel agent monotherapy	26	32.5%	5	4.1%
PI3K	5		4	
Bispecific antibody	4		0	
Checkpoint inhibitor	3		0	
BTKi	3		1	
CD19	2		0	
CD22	2		0	
Other N-of-1	7		0	
Other	7	8.8%	6	4.9%
Chemotherapy without CD20	1		6	
Checkpoint inhibitor based combinations	3		0	
Other N-of-1	3		0	

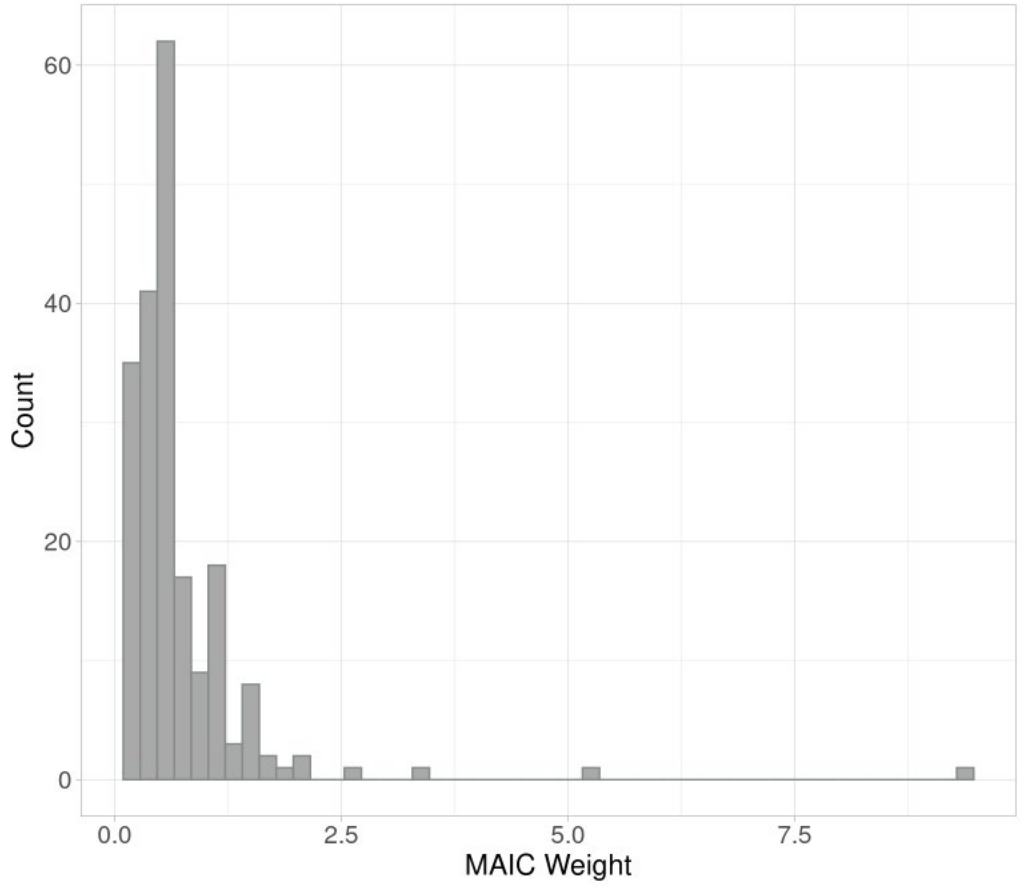
3) **Table S5.** Outcomes in Primary Analysis by Receipt of Index Therapy on a Clinical Trial

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

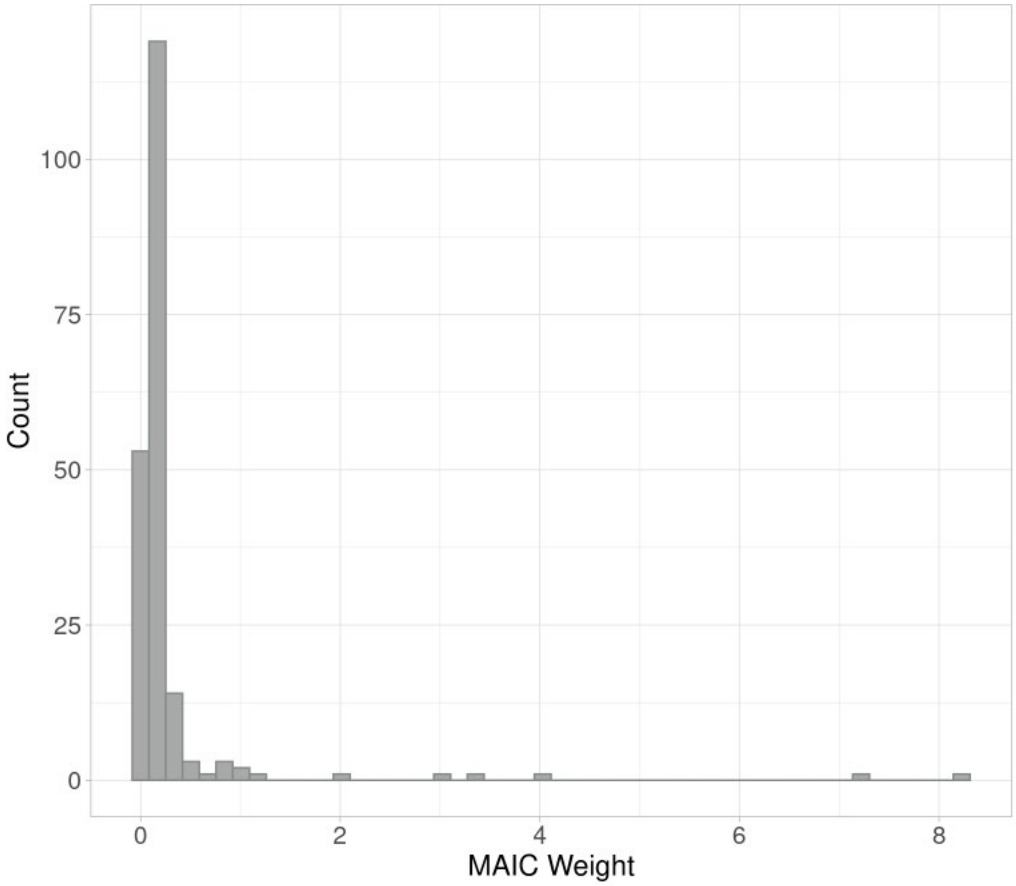
4) TTNT

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).



Lymphoma Epidemiology of Outcomes

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 3rd 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 publication[12].

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

Sample Size	True ORR	Clopper-Pearson 95% CI Limits	True ORR	Clopper-Pearson 95% CI Limits	True ORR	Clopper-Pearson 95% CI Limits
75	60%	48%, 71%	70%	58%, 80%	80%	69%, 88%
100	60%	50%, 70%	70%	60%, 79%	80%	71%, 87%
200	60%	53%, 69%	70%	63%, 76%	80%	74%, 85%
300	60%	54%, 66%	70%	65%, 75%	80%	75%, 84%

400	60%	55%, 65%	70%	65%, 75%	80%	76%, 84%
500	60%	56%, 64%	70%	66%, 74%	80%	76%, 83%

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 publication[12]) 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 publication[12], unless otherwise specified.

2.9.1 Matching-adjusted indirect comparison (MAIC)[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[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[12], 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

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. *CA Cancer J Clin.* 2015;65(1):5-29.
2. Tilly H, Gomes da Silva M, Vitolo U, et al. Diffuse large B-cell lymphoma (DLBCL): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* 2015;26 Suppl 5:v116-125.
3. Ambinder AJ, Shenoy PJ, Malik N, Maggioncalda A, Nastoupil LJ, Flowers CR. Exploring risk factors for follicular lymphoma. *Adv Hematol.* 2012;2012:626035.
4. 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.
5. Marcus R, Imrie K, Belch A, et al. CVP chemotherapy plus rituximab compared with CVP as first-line treatment for advanced follicular lymphoma. *Blood.* 2005;105(4):1417-1423.
6. Hiddemann W, Kneba M, Dreyling M, et al. Frontline therapy with rituximab added to the combination of cyclophosphamide, doxorubicin, vincristine, and prednisone (CHOP) significantly improves the outcome for patients with advanced-stage follicular lymphoma compared with therapy with CHOP alone: results of a prospective randomized study of the German Low-Grade Lymphoma Study Group. *Blood.* 2005;106(12):3725-3732.
7. Casulo C, Day B, Dawson KL, et al. Disease characteristics, treatment patterns, and outcomes of follicular lymphoma in patients 40 years of age and younger: an analysis from the National Lymphocare Studydagger. *Ann Oncol.* 2015;26(11):2311-2317.
8. Casulo C, Nastoupil L, Fowler NH, Friedberg JW, Flowers CR. Unmet needs in the first-line treatment of follicular lymphoma. *Ann Oncol.* 2017;28(9):2094-2106.
9. Rivas-Delgado A, Magnano L, Moreno-Velazquez 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.
10. 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.
11. 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. Published 2020 Jul 17. doi:10.1038/s41408-020-00340-z
12. Casulo C, Larson MC, Lunde MA, et al. Treatment Patterns and Outcomes of Patients with Relapsed/Refractory Follicular Lymphoma Receiving Three or More Lines of Systemic Therapy: Results from a Lymphoma Epidemiology of Outcomes Consortium Observational Study. In press, Lancet Haematology.
13. Signorovitch JE, Sikirica V, Erder MH, et al. Matching-Adjusted Indirect Comparisons: A New Tool for Timely Comparative Effectiveness Research. Value in Health, Volume 15, Issue 6, 2012, Pages 940-947, ISSN 1098-3015, https://doi.org/10.1016/j.jval.2012.05.004
14. Phillipo D, Ades A, Dias S, Palmer S, Abrams K and Welton N. Methods for population-adjusted indirect comparisons in health technology appraisal(2018). Medical Decision Making. 38(2): 200-211.

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