

ISSN: (Print) (Online) Journal homepage: <https://www.tandfonline.com/loi/iery20>

A 24-month updated analysis of the comparative effectiveness of ZUMA-5 (axi-cel) vs. SCHOLAR-5 external control in relapsed/refractory follicular lymphoma

M Lia Palomba, Paola Ghione, Anik R Patel, Myrna Nahas, Sara Beygi, Anthony J Hatswell, Steve Kanters, Eve H. Limbrick-Oldfield, Sally W Wade, Markqayne D Ray, Jessica Owen, Sattva S Neelapu, John Gribben, John Radford & Sabela Bobillo

To cite this article: M Lia Palomba, Paola Ghione, Anik R Patel, Myrna Nahas, Sara Beygi, Anthony J Hatswell, Steve Kanters, Eve H. Limbrick-Oldfield, Sally W Wade, Markqayne D Ray, Jessica Owen, Sattva S Neelapu, John Gribben, John Radford & Sabela Bobillo (2023) A 24-month updated analysis of the comparative effectiveness of ZUMA-5 (axi-cel) vs. SCHOLAR-5 external control in relapsed/refractory follicular lymphoma, Expert Review of Anticancer Therapy, 23:2, 199-206, DOI: [10.1080/14737140.2023.2171994](https://doi.org/10.1080/14737140.2023.2171994)

To link to this article: <https://doi.org/10.1080/14737140.2023.2171994>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 10 Feb 2023.



[Submit your article to this journal](#)



Article views: 605

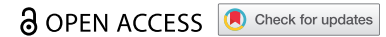


[View related articles](#)





[View Crossmark data](#)

ORIGINAL RESEARCH



A 24-month updated analysis of the comparative effectiveness of ZUMA-5 (axi-cel) vs. SCHOLAR-5 external control in relapsed/refractory follicular lymphoma

M Lia Palomba ^a, Paola Ghione^a, Anik R Patel^b, Myrna Nahas^b, Sara Beygi^b, Anthony J Hatswell^c, Steve Kanter^d, Eve H. Limbrick-Oldfield ^d, Sally W Wade^e, Markqayne D Ray^b, Jessica Owen^c, Sattva S Neelapuf^f, John Gribben^g, John Radford^h and Sabela Bobilloⁱ

^aDepartment of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA; ^bKite, A Gilead Company, Santa Monica, CA, USA; ^cDelta Hat, Nottingham, UK; ^dRainCity Analytics, Vancouver, BC, Canada; ^eWade Outcomes Research and Consulting, Salt Lake City, UT, USA; ^fThe University of Texas MD Anderson Cancer Center, Houston, TX, USA; ^gCancer Research UK Barts Centre, London, UK; ^hThe University of Manchester and Christie NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK; ⁱVall D'Hebron Institute of Oncology, Barcelona, Spain

ABSTRACT

Background: In the ZUMA-5 trial (Clinical trials identification: NCT03105336), axicabtagene ciloleucel (axi-cel; a chimeric antigen receptor T-cell therapy) demonstrated high rates of durable response in relapsed/refractory (r/r) follicular lymphoma (FL) patients and clear superiority relative to the SCHOLAR-5 external control cohort. We update this comparison using the ZUMA-5 24-month data.

Research design and methods: The SCHOLAR-5 cohort is comprised of r/r FL patients who initiated $\geq 3^{\text{rd}}$ line of therapy after July 2014 and meeting ZUMA-5 eligibility criteria. Groups were balanced for patient characteristics through propensity scoring on prespecified prognostic factors using standardized mortality ratio (SMR) weighting. The overall response rate was compared using a weighted logistic regression. Time-to-event outcomes were evaluated using a Cox regression.

Results: For SCHOLAR-5, the sum of weights for the 143 patients was 85 after SMR weighting, versus 86 patients in ZUMA-5. The median follow-up was 29.4 months and 25.4 months for ZUMA-5 and SCHOLAR-5, respectively. The hazard ratios for overall survival and progression-free survival were 0.52 (95% confidence interval (CI): 0.28–0.95) and 0.28 (95% CI: 0.17–0.45), favoring axi-cel.

Conclusion: This updated analysis, using a longer minimum follow-up than a previously published analysis, shows that the improved efficacy of axi-cel, relative to available therapies, in r/r FL is durable.

ARTICLE HISTORY

Received 1 December 2022
Accepted 11 January 2023

KEYWORDS



Follicular lymphoma;
comparative effectiveness;
axicabtagene ciloleucel;
propensity score analysis;
ZUMA-5


1. Introduction

Follicular lymphoma (FL) is the most common indolent [1] non-Hodgkin lymphoma (NHL) [2,3]. Whilst the prognosis of patients who respond to first-line treatment is good, relapsed/refractory (r/r) FL is generally considered incurable [4]. There is a growing body of real-world evidence in three lines or more ($\geq 3^{\text{rd}}$ line of treatment, LoT) r/r FL patients that improves our understanding of clinical outcomes of systemic treatments in general. For example, the rate of survival and durability of response both decrease with each subsequent LoT among patients who fail multiple LoTs [5–7]. The majority of these studies report a median progression-free survival (PFS) [5,6] or event-free survival (EFS) [8] of 11 ± 2 months, whilst the LEO-CReWE cohort had a median PFS of 17 months [9]. These studies also highlight a high degree of heterogeneity in the choice of treatments for r/r FL patients beyond the second LoT [6,9]. This heterogeneity may be due to a heretofore lack of a clearly superior treatment option and/or definitive clinical guidelines for higher lines for FL patients. Taken together, these recent studies suggest that there remains an unmet need for this clinically challenging population.

More recently, CD19-directed chimeric antigen receptor (CAR) T-cell therapies have emerged as a potentially more durable option for r/r FL patients. Both axicabtagene ciloleucel (axi-cel) and tisagenlecleucel (tisa-cel) were granted approvals for r/r FL patients who have failed two prior LoTs by the FDA in 2021 and 2022, respectively, [10,11], and by the European Medicines Agency in 2022 [12,13]. These approvals were based on single-arm phase II clinical trials [14,15], with ZUMA-5 – the axi-cel trial in patients with r/r FL – reporting a 94% overall response rate (ORR). In addition to ORR, ZUMA-5 has demonstrated a striking complete response (CR, 79%), 18-month overall survival (OS) rate (88%), and 18-month progression-free survival (PFS) rate (69%) in patients with r/r FL.

In the early stages of approval, non-comparative trials are crucial for demonstrating efficacy; however, the lack of a control arm precludes conclusions from being drawn about where axi-cel fits within the current therapeutic landscape. For ZUMA-5, an external control cohort named SCHOLAR-5 was designed to help draw such inferences. SCHOLAR-5 is a multi-country cohort that was previously compared with ZUMA-5 using its 18-month minimum follow-up data. The results of that analysis, in which SCHOLAR-5 individual patient data were

CONTACT M Lia Palomba  palombam@mskcc.org  Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/14737140.2023.2171994>

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

weighted using propensity score methods to match the ZUMA-5 patient characteristics and optimize internal validity, demonstrated that axi-cel offered a benefit in all measured efficacy outcomes.

Given FL's indolent nature, it is important to understand the comparative efficacy of axi-cel relative to other treatments over the long term. In this study, we sought to update the comparative analysis of ZUMA-5 to SCHOLAR-5, using the 24-month minimum follow-up data for ZUMA-5 in order to test if the axi-cel benefits are maintained at 2 years.

2. Patients and methods

2.1. Design and setting

This is an updated analysis of the comparative analysis of ZUMA-5 to SCHOLAR-5 using the 24-month follow-up ZUMA-5 data. Details of the study's design and analysis have previously been reported [1]. SCHOLAR-5 is comprised of two patient sub-cohorts. Sub-cohort A was constructed from patient records from seven institutions in five countries, extracted from electronic medical records dating from 2014 to 2020, with additional manual extraction from paper sources as required. Institutions would be eligible for SCHOLAR-5 if they were able to provide a minimum of 10 patients and meet criteria to ensure data quality and completeness. To minimize unobserved confounding related to treatment center characteristics, sites were limited to those that were similar to the ZUMA-5 trial sites (i.e. primarily high-volume tertiary academic centers). Sub-cohort B was included to supplement the real-world data. This cohort included patients who had taken part in the DELTA trial of idelalisib, a PI3K kinase (PI3K) inhibitor, that met the SCHOLAR-5 inclusion criteria. To avoid overrepresenting PI3K inhibitor treatments, only the first subsequent LoT after idelalisib of Subcohort B was included in this analysis. This selection better represented the heterogeneous mix of treatment options available to third line plus r/r FL patients. These data were collected using the DELTA trial case report form described in greater detail elsewhere [16]. Institutional Review Board approval for the study was obtained separately for each participating site. A detailed description of the SCHOLAR-5 methodology has previously been reported [1], and details of the ZUMA-5 trial cohort are reported elsewhere [17]. Investigators abided by the general ethical principles outlined in the Declaration of Helsinki and institutional review board approval for the study was obtained separately for each participating site.

The inclusion criteria for SCHOLAR-5 mirrored that of ZUMA-5, and included being an adult (≥ 18 years of age) diagnosed with r/r FL grade 1–3a, and initiating a third LoT or higher. Treatments involving anti-CD20 monotherapy, radiotherapy monotherapy, or surgery on its own were not eligible LoTs. In addition, patients were not included in the analysis if they had previously received CAR T-cell therapy or genetically modified therapy, had an Eastern Cooperative Oncology Group (ECOG) performance score > 1 , had transformed diffuse large B cell lymphoma or 3b FL histology, and had an index date before July 2014 or less than 12 months before the database cutoff date (i.e. patients were required to have the possibility of at least 12 months of follow-up for

inclusion). If a patient had at least one eligible LoT, they were included in the analysis.

Patients treated with axi-cel in ZUMA-5 were compared to those treated with other available treatment options in SCHOLAR-5. These included approved and experimental drug therapies, and autologous and allogeneic transplants.

2.2. Variables assessed

Prior to any analysis, variables that would be included in the propensity score model (methods described below) were pre-specified by the investigator team and external experts based on likely clinical relevance and prognostic value, in an aim to reduce the differences between the groups on these important variables [18]. These included progression of disease within 24 months of initiation first-line anti-CD20 combination therapy (POD24) status, number of prior LoT, relapsed vs. refractory to last LoT, prior stem cell transplant (SCT), tumor bulk (diameter of largest lesion), time from last treatment, best response to previous line, age, and prior exposure to anti-CD20 alkylator combination therapy. Where possible, missing data for these variables were supplemented using multiple imputations to enable inclusion of the most prognostic baseline variables in the propensity score specification. Multiple imputations were applied for variables with $< 40\%$ missing data in either dataset [19]. Imputation was chosen over a complete case analysis to avoid selection bias [20]. Retrospectively collected data often have missing values and are likely missing at random. In order to ensure the robustness of our primary analytical approach, we included a sensitivity analysis with only complete cases for variables used in the propensity score model. Further details on the variables included in the model, and handling of missingness (including the imputation of partially missing data, including for effectiveness outcomes) are described in the previous work [1].

In SCHOLAR-5, although not included in the propensity score model, if provided, the Karnofsky performance score was used to derive missing ECOG performance scores. In this cohort, the index date chosen was the initiation date of the index LoT. For sub-cohort-A, in cases where patients had multiple eligible LoTs, the index LoT was randomly chosen. This was considered an unbiased approach, which would likely lead to good overlap with the ZUMA-5 LoT distribution [21,22]. To reduce time-period bias due to the introduction of PI3K-inhibitors and other treatments, and because the Lugano criteria for disease assessment was formalized in 2014 [23], index date must have occurred after July 2014.

For all SCHOLAR-5 patients, CAR-T or any other cellular therapy were ineligible index treatments, and patients were censored if they received these treatments during follow-up. The index treatment line selection period extended from July 2014 to site-specific dates of abstraction, with the latest date being December 2020.

2.3. Endpoints

PFS, OS, time to next treatment (TTNT; including death as an event), duration of response (DOR), ORR, and CR were

included to assess ZUMA-5 compared to SCHOLAR-5. Methods of disease response and progression assessment did vary by cohort. In addition to the Lugano criteria, response assessments in sub-cohorts A and B included computed tomography (CT) scans using older criteria. In ZUMA-5, tumor response and progression were evaluated using positron emission tomography (PET)-diagnostic CT scans using Lugano criteria. Progression data were not collected for the subsequent LoT in the DELTA trial, therefore sub-cohort-B was not included in the PFS analysis.

2.4. Statistical methods

Propensity score methods, specifically standardized mortality ratio (SMR) weighting, were applied to account for the imbalance of baseline characteristics, which could be confounders when comparing ZUMA-5 to SCHOLAR-5. The SMR weighting allowed for the creation of an external comparator arm from SCHOLAR-5 with a distribution of covariates that resembled those in ZUMA-5. Standardized mean differences were computed and required to be less than 0.1 after weighting, which is considered a strict threshold. Details about model specifications were made without knowledge of how those decisions impacted effect estimates. Propensity score methods [18] and the resulting distribution of weights used in all analyses have been previously described [1].

A two-sided 95% confidence interval (CI) was used for all results, and all tests were performed on the 5% alpha level (two-sided). Continuous variables were assessed using weighted linear regression modeling to test for differences between ZUMA-5 and SCHOLAR-5, whereas categorical variables were compared using weighted logistic regression models. For time-to-event variables, the relative difference in hazard of the outcome between groups was estimated using a weighted Cox proportional hazard regression.

Additional subgroup and sensitivity analyses – including removal of the DELTA trial patients from SCHOLAR-5 and analyses without multiple imputation (i.e. complete case analysis) – were also conducted. For each subgroup and sensitivity analysis, propensity scoring methods were re-applied to ensure balance in covariates was maintained. Analyses were performed using R Software version 3.6.3.

3. Results

3.1. Patient characteristics

With the exception of follow-up time in ZUMA-5 patients, the patient characteristics remain as previously reported [1]. The SCHOLAR-5 cohort of 143 r/r FL patients was reduced to an effective sample size of 85 patients after applying the SMR weights derived through propensity score methods. For the ZUMA-5 patients, the data cut was 6 months after the 18-month data cut, so the same 86 patients previously reported met the current inclusion criteria of a minimum follow-up time of 24 months. The median follow-up time after index treatment was 29.4 months for ZUMA-5 and 25.4 months for SCHOLAR-5 after SMR weighting. Table 1 lists the baseline characteristics of ZUMA-5 and SCHOLAR-5 before and after

the SMR weighting, and a complete list of baseline variables is in supplemental Tables S1-S2. Prior to weighting, ZUMA-5 patients appeared to have a higher proportion of high-risk baseline characteristics than SCHOLAR-5, including POD24, median number of prior LoTs, and refractory to prior line, although median age and ORR to prior LoT were slightly higher in SCHOLAR-5. All variables included in the propensity score model were balanced (Standardized Mean Difference <0.1) after SMR weighting, including POD24, number of prior LoT, relapsed vs. refractory disease, prior SCT, size of largest node, response to prior LoT, time since last therapy and age. Both the FLIPI and disease stage were missing to such a degree in SCHOLAR-5 that we were unable to reliably assess any potential imbalance in these variables. As expected, given the lack of a standard of care of r/r FL, there was a wide range of regimens in the index LoT for SCHOLAR-5 [1].

3.2. Time to event outcomes

Time-to-event outcomes are summarized using Kaplan-Meier plots in Figure 1. At 24 months, 63.4% of patients in ZUMA-5 had neither progressed nor died, compared to 15% of SCHOLAR-5 patients (Table 2). Median PFS was 39.6 months in ZUMA-5, compared to 12.7 in SCHOLAR-5, with a hazard ratio (HR) of 0.28 ($p < 0.001$) showing a substantial reduction in the risk of progression in ZUMA-5 compared to SCHOLAR-5. Among the patients with ≥ 3 prior LoT, the 24-month PFS in ZUMA-5 patients was 59.0 (95% CI: 44.5–71.0) and in SCHOLAR-5 patients was 5.7% (95% CI: 0.0–12.2, $p < 0.001$) (Supplementary Figure S2), a result which held when the 5 ZUMA-5 patients without biopsy confirmed lack of transformation were excluded (Supplementary Table S3, Figure S3). As progression dates were not collected for the DELTA trial patients, the effective sample size for SCHOLAR-5 was 56 for PFS, but the weighted sample excluding DELTA trial patients was well matched to ZUMA-5 (Supplementary Table S1). The substantial reduction in the risk of progression held across all additional subgroup population analyses, including POD24, refractory patients, and patients with prior SCT (Supplementary Figures S4-S7).

OS at 24 months was 81.2% in ZUMA-5 and 63.4% in SCHOLAR-5. Median OS was not reached in ZUMA-5 and was 59.8 months in SCHOLAR-5, with a HR of 0.52 (95% CI: 0.28–0.95, $p < 0.001$), a 48% reduction in the risk of death (Table 2). Notably among patients with ≥ 3 prior LoT, OS improvements were more pronounced (HR: 0.43; 95% CI: 0.23–0.82, $p < 0.05$), with a 57% reduction in the risk of death, and a 24-month survival rate of 79.8% [95% CI: 67.1–88.0] and 51.5% [95% CI: 36.2–66.8] in ZUMA-5 and SCHOLAR-5, respectively. Findings were maintained across almost all pre-specified sensitivity analyses, including those with the DELTA trial cohort removed (Supplementary Figures S4-S7), highlighting the robustness of the data. Most sensitivity analyses led to a reduced HR, and therefore a larger treatment effect. This included the propensity score matching analysis (HR: 0.45; 95% CI: 0.21–0.93, $p < 0.05$), the complete case analysis without multiple imputation (HR: 0.47; 95% CI: 0.26–0.84, $p < 0.05$), and the safety cohort

Table 1. Demographic and clinical characteristics of patients in the comparative analysis set.

	SCHOLAR-5 before weighting (n = 143)	ZUMA-5 (n = 86)	Weighted SCHOLAR-5 (n = 85)	Weighted SMD
Median age* (range)	64 (36–89)	62 (34–79)	61 (36–89)	0.036
Male – no. (%)	81 (56.6%)	48 (55.8%)	53 (61.9%)	0.123
Median size of largest nodal mass* (IQR) – cm	4.2 (2.8–6.5)	4.4 (3.3–6.4)	4.0 (2.9–6.3)	0.094
Follicular lymphoma subtype – no. (%)				
Grade 1	56 (42.4)	20 (23.3)	30 (37.3)	0.54
Grade 2	61 (46.2)	43 (50)	42 (52.6)	
Grade 3a	15 (11.4)	23 (26.7)	8 (10.1)	
Missing	11	0	5	
Median number of prior lines of therapy (range)	2 (2–8)	3 (2–9)	3 (2–8)	0.047
Median time since last treatment* (IQR) – months	6.8 (1.2–22.7)	3.5 (1.8–9.0)	2.3 (0.7–8.0)	0.056
Response to prior line of therapy* – no (%)				
CR	41 (28.7)	23.01 (26.8)	19 (22.8)	0.073
PR	49 (34.3)	19.34 (22.5)	19 (22.4)	
SD	22 (15.4)	24.15 (28.1)	26 (31.2)	
PD	31 (21.7)	19.5 (22.7)	20 (23.5)	
Refractory to prior LoT* – no. (%)	87 (60.6)	63 (73.3)	65 (76.6)	0.077
POD24* – no. (%)	51 (35.7)	49 (57.0)	47 (55.9)	0.022
Prior stem cell transplant* – no. (%)	31 (21.7)	21 (24.4)	24 (28.0)	0.08
Median time since diagnosis (IQR) – months	84.8 (53.0–130.5)	59.9 (35.1–96.6)	64.6 (41.0–115.8)	0.10
Disease stage – no. (%)				
I	4 (6.2)	2 (2.3)	1 (4.6)	NE
II	2 (3.1)	9 (10.5)	0 (1.3)	
III	17 (26.2)	35 (40.7)	8 (27.0)	
IV	42 (64.6)	40 (46.5)	20 (67.1)	
Missing	78	0	55	
Number of nodal sites – no. (%)				
1	14 (15.1)	16 (22.5)	8 (14.1)	NE
2	17 (18.3)	12 (16.9)	13 (21.5)	
3	9 (9.7)	7 (9.9)	7 (10.9)	
>4	53 (57)	36 (50.7)	32 (53.6)	
Missing	50	15	25 (29.4)	
FLIPI – no. (%)				
0	2 (4)	3 (3.5)	0 (0.4)	NE
1	4 (8)	10 (11.6)	2 (9.5)	
2	11 (22)	33 (38.4)	4 (17.4)	
3	19 (38)	25 (29.1)	7 (32.8)	
4	10 (20)	12 (14.0)	6 (28.1)	
5	4 (8)	3 (3.5)	3 (11.7)	
Missing	93	0	62	

* Variables used in propensity score weighting. The SMD for disease stage, number of nodal sites, and FLIPI were not evaluable due to missing data. See supplemental Table S5 for the SMD values before weighting. CR, complete response; FLIPI, Follicular Lymphoma International Prognostic Index; IQR, inter-quartile range; LoT, line of therapy; NE, not evaluable; PD, progressive disease; POD24, progression of disease within 24 months of starting first line chemo-immunotherapy; PR, partial response; SD, stable disease, SMD; standardized mean difference. This table was originally published in *Blood*, the journal of the American Society of Hematology (ASH).

that included all ZUMA-5 patients (HR: 0.48; 95% CI: 0.26–0.89, $p < 0.05$). These improvements were also seen in other endpoints including PFS and response rates.

Median TTNT was 39.6 months in ZUMA-compared to 23.4 months in SCHOLAR-5, with a 42% reduction in the risk of having initiated a new line of treatment (HR: 0.58; 95% CI:

Table 2. Comparison of outcomes between SCHOLAR-5 and ZUMA-5.

	SCHOLAR-5	ZUMA-5	Treatment effect (95% CI)
Response outcomes			
ORR	42 (49.9%)	81 (94.2%)	OR: 16.2 (5.6–46.9) $p < 0.001$
CR	25 (29.9%)*	68 (79.1%)**	OR: 8.9 (4.3–18.3) $p < 0.001$
Time-to-event outcomes			
PFS			
Median (95% CI)	12.7 (6.2–4.7)	39.6 (25.7–NE)	HR: 0.28 (0.18–0.45) $p < 0.001$
24 months % (95% CI)	15.0 (4.8–25.2)	63.4 (51.6–73)	
OS			
Median (95% CI)	59.8 (21.9–NE)	NR (39.6–NE)	HR: 0.52 (0.28, 0.95) $p < 0.05$
24 months % (95% CI)	63.4 (50.3–76.4)	81.2 (71.2–88.1)	
TTNT			
Median (95% CI)	23.4 (9.5–NE)	39.6 (28.0–NE)	HR: 0.58 (0.36–0.95) $p < 0.05$
24 months % (95% CI)	49.5 (36.3–62.7)	63.8 (52.7–073)	

Note that rounding of patients after classifying as responders or non-responders in the SCHOLAR-5 weighted sample may lead to a small variability in total sample size.

CR, complete response; HR, hazard ratio; OR, odds ratio; ORR, overall response rate; OS, overall survival; TTNT, time-to-next treatment; CI, confidence interval.

* Response assessments includes CT-based and PET-Based scans with limited confirmatory bone marrow biopsies

**13 patients with imaging CRs did not receive a confirmatory bone marrow biopsy

0.36, 0.95, $p < 0.05$). Importantly, the assessment of TTNT was comparable between patient groups as initiation data of the next treatment is routinely captured as part of routine clinical care. Duration of response is presented in Supplementary **Figure S1**.

3.3. Response outcomes

ORR and CR remained unchanged from the 18-month analysis and were substantially higher in ZUMA-5 (ORR: 94%; CR: 79%) than SCHOLAR-5 (ORR: 50%; CR: 30%), with odds ratios of 16.2 ($p < 0.001$) and 8.9 ($p < 0.001$), respectively (Table 2).

4. Discussion

At 24 months, this study provides the longest available follow-up for a comparison between CAR-T and an external control cohort in r/r FL. The ZUMA-5 trial, with a minimum follow-up of 24-months, was compared to the SCHOLAR-5 external control cohort, which is comprised of patients from seven international cancer centers and post-trial patients from the DELTA trial. In order to minimize the issues of a non-randomized study design, propensity score methods were used to align the treatment groups with respect to effect-modifiers and prognostic factors. The efficacy benefit of axi-cel relative to the standard of care that was observed in the 18-month analysis was maintained at 24 months, suggesting that the treatment effect of axi-cel is durable. As the SCHOLAR-5 comparator remained the same, any changes in the comparison are solely due to events that occurred in the 6-month additional follow-up in ZUMA-5.

As with the previous study, the analytical methods used in this study represent best practice to optimize internal validity in a non-randomized study design [22]. This 24-month analysis used the same SMR weights as the 18-month analysis –

meaning that the alignment of baseline characteristics between ZUMA-5 and SCHOLAR-5 were as closely balanced statistically and clinically as they were previously [1]. The results were similar to those from the previous analysis. For PFS, the hazard ratio shifted from 0.30 at 18-months to 0.28 at 24-months and remained convincingly statistically significant ($p < 0.001$). The OS estimated hazard ratio also remained statistically significant, shifting from 0.42 to 0.52 ($p < 0.05$). Similarly, axi-cel continued to show superiority over the remaining time-to-event variable, TTNT ($p < 0.05$). Unsurprisingly, there were no changes in the responders and non-responders from this study between 18-months and 24-month follow-up. Altogether, axi-cel continues to show a striking improvement compared to previously available therapies.

Randomized controlled trials (RCTs) remain the gold standard with respect to epidemiological study designs and the best studies to assist with decision-making [24]. By design, RCTs optimize internal validity through the alignment of all prognostic factors and effect modifiers – whether they are observed. On the other hand, propensity score methods can only align with the observed covariates and remain at risk of both confounding and selection bias. However, in patient populations with substantially high unmet needs, such as this one, the utility of external cohort matched comparative studies is crucial to answer critical questions faster, and ultimately gain approvals and guide clinical decision-making prior to data from an RCT becoming available [25]. Whilst the findings of this study suggest that a promising clinical benefit may be associated with axi-cel in r/r FL, this should be studied further in future comparative studies. In the CAR T-cell therapeutic space, single-arm trials are much more common than their randomized counterparts. Nonetheless, in large B-cell lymphoma, results from three-phase III RCTs were recently published, years after non-

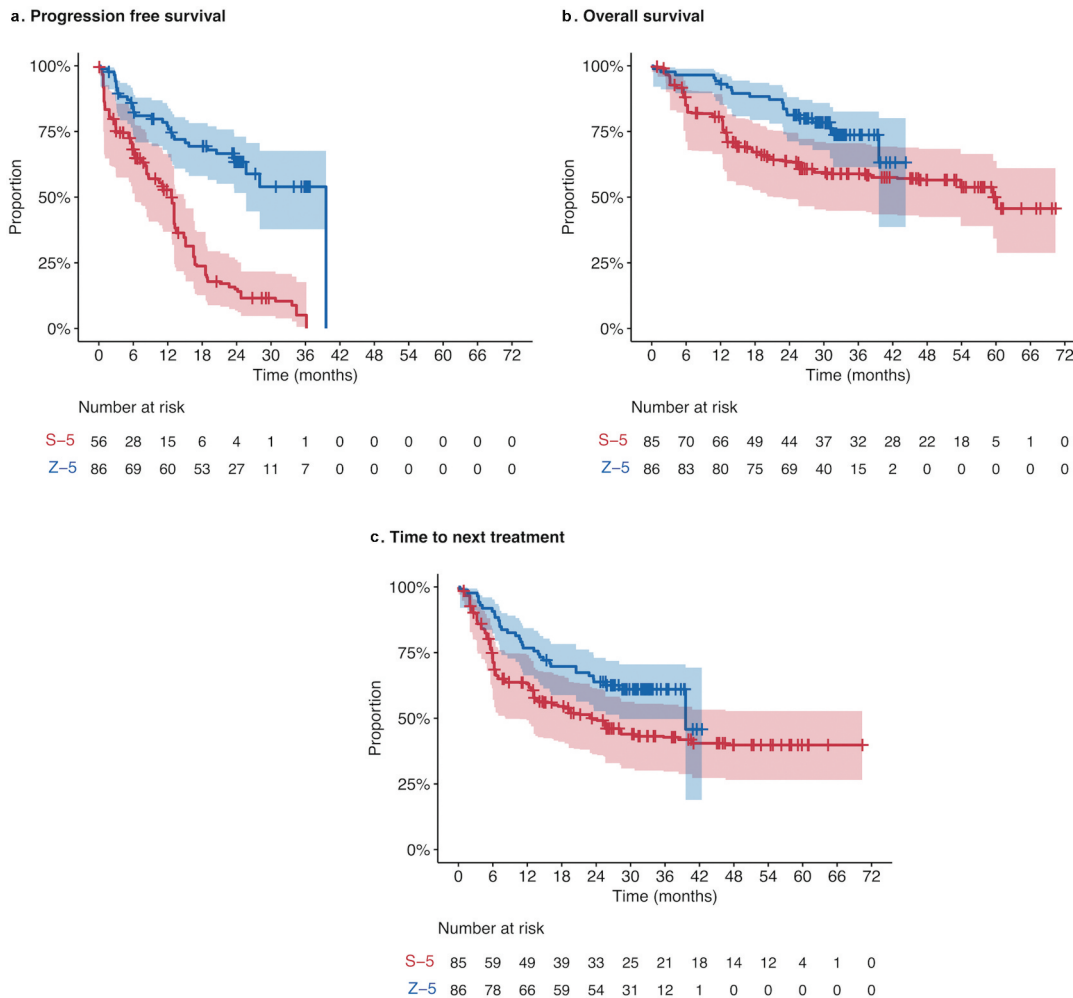


Figure 1. Kaplan-Meier plots comparing ZUMA-5 to SCHOLAR-5 for A. Progression-free survival; B. Overall survival; and C. Time-to-next-treatments.

Kaplan-Meier curves showing A. progression-free survival, B. overall survival, and C. time-to-next treatment in ZUMA-5 (blue), compared to SCHOLAR-5 (red). The shaded area represents 95% confidence interval. The number at risk for the SCHOLAR-5 analysis of PFS was reduced, due to the exclusion of DELTA participants from this analysis, as PFS was not available in this subgroup. See supplemental Figure S2 for results of all time to event outcomes with DELTA participants excluded prior to SMR weight

comparative phase II trials were used to obtain initial regulatory approvals [26–28]. To this end, it is reasonable to expect randomized studies to be presented in regulatory submissions in the future.

The SCHOLAR-5 population was selected to match the ZUMA-5 population, to obtain the degree of overlap required between the two populations to make a propensity score analysis feasible. Using an external control cohort in this way has shown results and effect sizes in large B-cell lymphoma populations [29] that are being replicated in real-world results of CAR-T compared to standard-of-care [30,31]. Results using this methodology are also beginning to be reported using external control cohorts to estimate the relative efficacy of tisa-cel to previously available therapies. The results from ELARA, the single-arm trial of tisa-cel in r/r FL patients who had failed two prior lines of treatment, have been compared to individual patient data from RECORD-FL (including a subgroup analysis using treatments post-2014) [32], showing that tisa-cel offers a meaningful benefit for patients over previously available treatments. As more comparisons with other available treatments emerge, it may be tempting to compare the treatment effects observed for axi-cel and other emerging therapies for r/r FL,

including tisa-cel and mosunetuzumab [33], using the real-world cohorts as a common comparator, and draw conclusions about the relative efficacy of the two treatments. Particularly, anchored networks are preferable to unanchored networks for indirect treatment comparisons [34]. Whilst there is value in this approach, offering the opportunity to establish relative efficacy of as yet un-compared treatments for r/r FL, this requires the thoughtful application of indirect-treatment comparison methodologies, mature data from both treatments, and the alignment of definitions that differ between the two trials, including POD24, treatments eligible as LoTs, and double refractory.

This study has potential limitations, which are all shared with the previous 18-month study. As expected in a retrospective study, some covariates of interest could not be included in the models due to missing values. These included FLIPI and bone marrow involvement. Multiple efforts were taken to minimize missingness and the resulting analyses including a large number of variables. Nonetheless, FLIPI was identified as a prognostic factor. Similarly, there may be confounding due to the unobserved variables – such as medical history.

In addition, there may have been measurement bias due to the differences in response assessment between the two groups of patients. These differences could both bias the results in favor of, and against, axi-cel. For example, ZUMA-5 used stringent criteria with central review and more frequent disease assessments than SCHOLAR-5. In addition, the response assessments were heterogeneous within SCHOLAR-5, compared to ZUMA-5, which used PET-CT scans. A CT assessment, in the absence of a PET assessment, could underestimate response rates in cases where a partial response was later reclassified as an overall disease assessment of CR based on a PET assessment. However, a PET scan alone could overestimate CR as compared to a CT scan alone. The exact direction of the impact of any measurement bias is not clear; however, this issue would likely not impact PFS or OS.

5. Conclusion

The strong, durable treatment effects of axi-cel demonstrated in our study suggest that CAR T-cell therapeutics could help resolve the unmet need in $\geq 3^{\text{rd}}$ LoT FL patients identified in recent studies. The durability of the results is particularly interesting given the observed decrease in durability after each passing line of existing treatments. Our previous study helped demonstrate that axi-cel provides clinically meaningful improvements over competing treatments with respect, including OS and PFS. This updated comparative analysis between ZUMA-5 and SCHOLAR-5 suggests that these results remain stable over time. Longer term studies, as well as randomized studies, such as the currently recruiting ZUMA-22 [35], will be needed to further understand the potential role of axi-cel for the treatment of the r/r FL population.

Acknowledgments

We thank the patients who participated in this study and their families, caregivers, and friends.

Funding

This manuscript was funded by Kite, a Gilead Company.

Declaration of interests

ML Palomba holds individual stocks and stock options from Seres and Notch, has received consulting fees from Novartis, Kite, PCYC, Cellectar, and BeiGene, research funding from Seres, patents and royalty fees from Seres, Juno, Wolters, and Kluwer; Paola Ghione reports no conflict of interest; AR Patel, M Nahas, and S Beygi are current employment at and holding stock and stock options from Kite, a Gilead company; AJ Hatswell is a current employment at Delta Hat; S Kanters and EH Limbrick-Oldfield are current employees at RainCity Analytics; SW Wade has received consulting fees from Kite, a Gilead company, Allergan, Johnson & Johnson; MD Ray is a current employee at and holds stock and stock options from Kite, a Gilead company; J Owen is a current employment at Delta Hat; SS Neelapu has received research support from Kite/Gilead, BMS, Cellectis, Poseida, Allogene, Precision Biosciences, and Adicet Bio; has served as Advisory Board Member/Consultant for Kite/Gilead, Merck, Novartis, Sellas Life Sciences, Athenex, Allogene, Incyte, Adicet Bio, BMS, Legend Biotech, Bluebird Bio, Fosun Kite, Sana Biotechnology, Caribou, Astellas Pharma, Morphosys, Janssen, Chimagen, ImmunoACT, and Orna Therapeutics; has received royalty income from Takeda Pharmaceuticals; has stock options from Longbow Immunotherapy, Inc; and has intellectual property related

to cell therapy.; J Gribben has received consulting fees from AbbVie, Acerta Group Limited/AstraZeneca, Bristol Myers Squibb/Celgene Corporation, Janssen Karyopharm, Morphosys AG, Novartis, and TG Therapeutics, research funding from Acerta Group Limited/AstraZeneca, Bristol Myers Squibb/Celgene Corporation, and Janssen. J Radford has received consulting fees from Takeda, ADCT, BMS, and Novartis, honoraria from Takeda, ADCT, and BMS, fees for serving on the speakers' bureau from Takeda and ADCT, research funding from Takeda, holding stocks or other ownerships on ADC Therapeutics and AstraZeneca; S Bobillo has received fees for serving on a speaker's bureau from Roche, Gilead, and Janssen, and consulting fees from Roche.

Author contributions statement

The study design and analysis were conducted in a collaboration between Kite, a Gilead Company (study sponsor), and the authors. M. Lia Palomba, Paola Ghione, Sabela Bobillo, Anik R. Patel, John Radford, and John Gribben contributed to data collection and verification. Steve Kanters, Eve Limbrick-Oldfield, Jessica Owen, and Anthony J. Hatswell contributed to data verification and data analysis. All authors contributed to the results interpretation, writing the manuscript and approved the final submitted version. The corresponding author had final responsibility for the decision to submit for publication.

Data sharing

All data are confidential. They can be made available upon approval of a research proposal and signed data access agreement.

ORCID

M Lia Palomba  <http://orcid.org/0000-0001-5099-9156>

Steve Kanters  <http://orcid.org/0000-0003-0322-5398>

Eve H. Limbrick-Oldfield  <http://orcid.org/0000-0001-5400-8030>

References

Papers of special note have been highlighted as either of interest (*) or of considerable interest () to readers.**

- Ghione P, Palomba ML, Patel AR, et al. Comparative effectiveness of ZUMA-5 (axi-cel) vs SCHOLAR-5 external control in relapsed/refractory follicular lymphoma. *Blood*. 2022 Aug 25;140(8):851–860.
- **This study is the 18-month analysis comparing ZUMA-5 to SCHOLAR-5 on which this builds from.**
- Al-Hamadani M, Habermann TM, Cerhan JR, et al. Non-Hodgkin lymphoma subtype distribution, geodemographic patterns, and survival in the US: a longitudinal analysis of the National Cancer Data Base from 1998 to 2011 [Research Support, Non-U.S. Gov't]. *Am J Hematol*. 2015 Sep;90(9):790–795. Epub 2015 Jul 27
- Teras LR, DeSantis CE, Cerhan JR, et al. 2016 US lymphoid malignancy statistics by World Health Organization subtypes. *CA Cancer J Clin*. 2016 Nov 12;66(6):443–459.
- Qualls D, Salles G. Prospects in the management of patients with follicular lymphoma beyond first-line therapy. *Haematologica*. 2022 Jan 1;107(1):19–34.
- 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 Jul 17;10(7):74.
- Ghione P, Ghesquieres H, Bobillo S, et al. Outcomes in later-lines of therapy for relapsed/refractory follicular lymphoma: results from the international SCHOLAR-5 study. *Hematol Oncol*. 2021;39(S2):S2.
- Kanters S, Kahl BS, Wiesinger A, et al. Clinical outcomes in patients relapsed/refractory after ≥ 2 prior lines of therapy for follicular lymphoma: a systematic literature review and meta-analysis. *J Clin Oncol*. 2021 May 20;39(15_suppl):e19548–e19548.
- Salles G, Schuster SJ, Fischer L, et al. A retrospective cohort study of treatment outcomes of adult patients with relapsed or refractory follicular lymphoma (ReCORD-FL). *Hemasphere*. 2022 Jul;6(7):e745.

9. Casulo C, Larson MC, Lunde JJ, et al. Treatment patterns and outcomes of patients with relapsed or refractory follicular lymphoma receiving three or more lines of systemic therapy (LEO CReWE): a multicentre cohort study. *Lancet Haematol.* 2022 Apr;9(4):e289–e300.
10. Food and Drug Administration. FDA grants accelerated approval to axicabtagene ciloleucel for relapsed or refractory follicular lymphoma: u.S. Food and Drug Administration; 2021a [2022 Sept 6th]. Available from: <https://www.fda.gov/drugs/resources-information-approved-drugs/fda-grants-accelerated-approval-axicabtagene-ciloleucel-relapsed-or-refractory-follicular-lymphoma>
11. Food and Drug Administration. FDA approves tisagenlecleucel for relapsed or refractory follicular lymphoma: u.S. Food and Drug Administration; 2021b [2022 Sept 6th]. Available from: <https://www.fda.gov/drugs/resources-information-approved-drugs/fda-approves-tisagenlecleucel-relapsed-or-refractory-follicular-lymphoma>
12. European Medicines Agency. New indication concerns the treatment of adult patients with relapsed or refractory follicular lymphoma 2022a [2022 Sept 6th]. Available from: <https://www.esmo.org/oncology-news/ema-recommends-extension-of-therapeutic-indications-for-axicabtagene-ciloleucel>
13. European Medicines Agency. New indication concerns the treatment of adult patients with relapsed or refractory follicular lymphoma after two or more lines of systemic therapy 2022b [2022 Sept 6th]. Available from: <https://www.esmo.org/oncology-news/ema-recommends-extension-of-therapeutic-indications-for-tisagenlecleucel>
14. Jacobson C, Chavez JC, Sehgal AR, et al. Primary Analysis of Zuma-5: a Phase 2 Study of Axicabtagene Ciloleucel (Axi-Cel) in Patients with Relapsed/Refractory (R/R) Indolent Non-Hodgkin Lymphoma (iNHL). *Blood.* 2020;136(Supplement 1):40–41.
15. Fowler NH, Dickinson M, Dreyling M, et al. Tisagenlecleucel in adult relapsed or refractory follicular lymphoma: the phase 2 ELARA trial. *Nat Med.* 2022 Feb;28(2):325–332.
16. Gopal AK, Kahl BS, de Vos S, et al. PI3K δ inhibition by idelalisib in patients with relapsed indolent lymphoma. *N Engl J Med.* 2014 March 13;370(11):1008–1018.
17. Jacobson CA, Chavez JC, Sehgal AR, et al. Axicabtagene ciloleucel in relapsed or refractory indolent non-Hodgkin lymphoma (ZUMA-5): a single-arm, multicentre, phase 2 trial. *Lancet Oncol.* 2021 Dec 8;22(12):1005–1018.
- This publication provides further details on the ZUMA-5 pivotal trial that is central to the current study.
18. Brookhart MA, Schneeweiss S, Rothman KJ, et al. Variable selection for propensity score models. *Am J Epidemiol.* 2006 Jun 15;163(12):1149–1156.
- This publication provides further details on the key methods used in our study to emulate randomized trials.
19. Jakobsen JC, Gluud C, Wetterslev J, et al. When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts. *BMC Med Res Methodol.* 2017 December 06;17(1). DOI:10.1186/s12874-017-0442-1.
20. Ali MS, Prieto-Alhambra D, Lopes LC, et al. Propensity score methods in health technology assessment: principles, extended applications, and recent advances [Review]. *Front Pharmacol.* 2019 September 18;10(973). 10.3389/fphar.2019.00973.
21. Hernán MARJ. Causal Inference: what If. Boca Raton: Chapman & Hall/CRC; 2020.
22. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol.* 2016 Apr 15;183(8):758–764.
23. Cheson BD, Fisher RI, Barrington SF, et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. *J Clin Oncol.* 2014 October 20;32(27):3059–3067.
24. Hariton E, Locascio JJ. Randomised controlled trials - the gold standard for effectiveness research: study design: randomised controlled trials. *Bjog.* 2018 Dec;125(13):1716.
25. Ribeiro TB, Buss L, Wayant C, et al. Comparison of FDA accelerated vs regular pathway approvals for lung cancer treatments between 2006 and 2018. *PLoS One.* 2020;15(7):e0236345.
26. Bishop MR, Dickinson M, Purtil D, et al. Second-line tisagenlecleucel or standard care in aggressive B-cell lymphoma. *N Engl J Med.* 2021;386(7):629–639.
27. Kamdar M, Solomon SR, Arnason J, et al. Lisocabtagene maraleucel versus standard of care with salvage chemotherapy followed by autologous stem cell transplantation as second-line treatment in patients with relapsed or refractory large B-cell lymphoma (TRANSFORM): results from an interim analysis of an open-label, randomised, phase 3 trial. *Lancet.* 2022 Jun 18;399(10343):2294–2308.
28. Locke FL, Miklos DB, Jacobson CA, et al. Axicabtagene ciloleucel as second-line therapy for large B-cell lymphoma. *N Engl J Med.* 2021;386(7):640–654.
29. Summary basis for regulatory action, (2017).
30. Jacobson CA. Real-world evidence of axicabtagene ciloleucel (Axi-cel) for the treatment of large B-cell lymphoma (LBCL) in the United States (US). *ASCO* 2021.
31. Nastoupil LJ, Jain MD, Feng L, et al. Standard-of-care axicabtagene ciloleucel for relapsed or refractory large B-cell lymphoma: results from the US lymphoma CAR T consortium. *J Clin Oncol.* 2020 Sep 20;38(27):3119–3128.
32. Salles G, Schuster SJ, Dreyling MH, et al. Efficacy comparison of tisagenlecleucel versus standard of care in patients with relapsed or refractory follicular lymphoma. *Blood.* 2021;138(Supplement 1):3528.
33. Budde LE, Sehn LH, Matasar M, et al. Safety and efficacy of mosunetuzumab, a bispecific antibody, in patients with relapsed or refractory follicular lymphoma: a single-arm, multicentre, phase 2 study. *Lancet Oncol.* 2022;23(8):1055–1065.
34. Phillippo DM, Ades AE, Dias S, et al. Methods for population-adjusted indirect comparisons in health technology appraisal. *Med Decis Making.* 2018 Feb;38(2):200–211.
35. Study of axicabtagene ciloleucel versus standard of care therapy in participants with relapsed/refractory follicular lymphoma (ZUMA-22) (NCT05371093) [cited 2022]. Available from: <https://clinicaltrials.gov/ct2/show/NCT05371093>