VOLUME 35 · NUMBER 5 · FEBRUARY 10, 2017

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Thirty-Month Complete Response as a Surrogate End Point in First-Line Follicular Lymphoma Therapy: An Individual Patient-Level Analysis of Multiple Randomized Trials

Qian Shi, Christopher R. Flowers, Wolfgang Hiddemann, Robert Marcus, Michael Herold, Anton Hagenbeek, Eva Kimby, Howard Hochster, Umberto Vitolo, Bruce A. Peterson, Emmanuel Gyan, Michele Ghielmini, Tina Nielsen, Sabine De Bedout, Tommy Fu, Nancy Valente, Nathan H. Fowler, Eva Hoster, Marco Ladetto, Franck Morschhauser, Emanuele Zucca, Gilles Salles, and Daniel J. Sargent†

Author affiliations and support information (if applicable) appear at the end of this article.

†Deceased

Published at ascopubs.org/journal/jco on December 27, 2016.

Written on behalf of the Follicular Lymphoma Analysis of Surrogacy Hypothesis (FLASH) group.

Corresponding author: Qian Shi, PhD, Cancer Center Statistics, Department of Health Science Research, Mayo Clinic, Harwick 8-28, 200 First St SW, Rochester, MN 55905; e-mail: shi.qian2@mayo.edu.

0732-183X/17/3505w-552w/$20.00

ABSTRACT

Purpose

Follicular lymphoma (FL) is an indolent cancer, with effective but rarely curative treatment options. As a standard study end point for first-line FL therapy, progression-free survival (PFS) requires extended follow-up (median PFS, > 7 years). To provide patients with earlier access to newer therapies, an earlier end point to expedite clinical trials is needed. Our objective was to formally assess the complete response rate at 30 months (CR30) after initiation of induction therapy as a potential surrogate end point for PFS in first-line FL therapy.

Patients and Methods

We analyzed individual patient data from 13 randomized multicenter trials of induction and maintenance regimens in first-line FL therapy published after 1990 and with sufficient data to evaluate whether CR30 could predict treatment effects on PFS. Correlation of the CR30 odds ratio with the PFS hazard ratio was evaluated by both linear regression (R^2_{WLS}) and bivariate copula (R^2_{Copula}) models. Prespecified criteria for surrogacy required either R^2_{WLS} or $R^2_{Copula} \geq 0.80$, with a lower-bound 95% CI > 0.60.

Results

Data from eight induction and five maintenance randomized trials in 3,837 evaluable patients were analyzed. The prespecified surrogacy threshold was met, with an R^2_{WLS} of 0.88 (95% CI, 0.77 to 0.96) and an R^2_{Copula} of 0.86 (95% CI, 0.72 to 1.00). Multiple sensitivity and supplemental analyses supported the robustness of the findings. A minimum 11% absolute improvement in CR30 from a 50% control rate predicted a significant treatment effect on PFS (hazard ratio, 0.69).

Conclusion

This large, prospective, pooled analysis of randomized chemotherapy, immunotherapy, and chemoimmunotherapy trials demonstrates that CR30 is a surrogate end point for PFS in first-line FL treatment trials. Use of this end point may expedite therapeutic development with the intent of bringing novel therapies to this patient population years before PFS results are mature.

J Clin Oncol 35:552-560. © 2016 by American Society of Clinical Oncology

INTRODUCTION

Follicular lymphoma is the most common indolent form of non-Hodgkin lymphoma, with an estimated 16,000 new cases[1] occurring annually in the United States. Although median survival now exceeds 10 years, follicular lymphoma remains largely incurable, characterized by repeated remission and recurrence over many years.

New treatment approaches remain a critical need. Progression-free survival (PFS) is presently the principal end point for regulatory approval of new agents. However, advances in treatment and the indolent nature of follicular lymphoma challenge the use of PFS as the primary end point in clinical trials, where median PFS now approaches 6 to 8 years.[2,3] This extended PFS prolongs trial duration, hinders the efficiency of meeting regulatory end points, lengthens exposure to ineffective treatments within trials, and limits the ability to expeditiously provide new effective therapeutic options. Therefore, the identification of alternative (ie, surrogate)

ASSOCIATED CONTENT

Appendix
DOI: 10.1200/JCO.2016.70.8651

DOI: 10.1200/JCO.2016.70.8651

end points that are measured earlier but can reliably predict PFS treatment effects is a critical need.

Overall response rate has been considered an end point for Food and Drug Administration accelerated approval, but it is most commonly used in late-line treatment settings or in high-risk populations.[4] In follicular lymphoma, complete response requires the disappearance of detectable clinical, biochemical, and radiographic evidence of disease and disappearance of all disease-related symptoms. Patients who achieve complete response have better outcomes than patients who achieve partial response or stable disease.[5-8] Results from randomized studies have suggested that in general, treatments with superior PFS have higher complete response rates.[3,9,10] An observational study showed that early progression predicted poor outcome,[11] and literature-based meta-analyses have identified a correlation between response rate and progression-related end points in indolent non-Hodgkin lymphoma.[12,13] However, as a result of variations in end point definitions and lack of individual patient data, trial-level correlation is difficult to interpret, and patient-level correlation cannot be assessed.

To formally assess complete response rate at a specific time point as a potential surrogate end point, the Follicular Lymphoma Analysis of Surrogate Hypothesis (FLASH) group conducted a prospectively planned pooled analysis of individual patient data from randomized controlled trials of first-line therapy in follicular lymphoma.

PATIENTS AND METHODS

Trial Selection

In September 2011, a comprehensive search of publications on follicular lymphoma in Medline maintained by the US National Library of Medicine was conducted with the use of the search terms follicular lymphoma AND randomized, indolent lymphoma AND randomized, and low-grade lymphoma AND randomized. Multiple cooperative, collaborative, and academic groups from North America and Europe also were approached for potentially completed and soon-to-be-published studies. Eligible studies were multicenter randomized trials in patients with previously untreated follicular lymphoma with a sufficient sample size ($\geq$ 100 in total or $\geq$ 50 cases of follicular lymphoma) and published after 1990. Trials that evaluated only induction treatment versus observation and studies without prespecified imaging, including computed tomography (CT), were excluded. All trials assessed complete response by using clinical and CT criteria.

The owners of all identified studies that met the prespecified inclusion eligibility criteria were contacted for individual patient data sharing. For the studies that transferred data, the feasibility of calculating 30-month complete response status was evaluated. Authors who refused data sharing or who did not provide sufficient data to determine 30-month complete response status were excluded. All exclusions were based on data quality and availability only and were determined before statistical analysis of the end points.

Surrogate End Point Candidates

On the basis of the published data and the balance between reducing the time required for trial conduct and accurately assessing the full treatment effect, the complete response rate at 30 months (CR30) after enrollment (ie, initiation of induction treatment) was prospectively selected as the principal surrogate candidate. This time point corresponds to the completion of current standard-of-care treatment, including 6-month induction and 2-year maintenance treatment.[1,11,14] Unconfirmed complete response was not considered for the primary analysis; response assessments required CT scan but not bone marrow evaluation, and investigator determination of complete response was accepted. The secondary surrogate end point candidate was complete response rate at 24 months (CR24) to be evaluated only if the surrogacy of CR30 was established.

The disease assessment measure at exactly 30 months after enrollment was not always available because of variations in specific time points for response evaluation and recording. When complete response (non-complete response) status was observed both before and after the 30-month time point, the end point status could be determined unambiguously. When response status observed before and after the 30-month time point differed, the determination of 30-month complete response status required calculation rules. For these cases, any assessment result within a window of 27 to 33 months (inclusive) after enrollment provided complete response status at 30 months. When no disease was observed within the window, 30-month complete response status could not be determined (ie, missing values). A predefined cut point of $>$ 20% of patients with missing complete response status data within a study after calculations resulted in study exclusion from surrogacy analyses.

Statistical Methods

True end point. The primary clinical end point was PFS, defined as time from enrollment to first documented disease progression or death as a result of any cause, whichever occurred first. Living patients without documented disease progression were censored on the date of last disease evaluation. Both primary end point and surrogate candidate were derived according to consistent calculation rules across studies.

General statistical methods. Within-trial treatment effects and 95% CIs for PFS and complete response end points were quantified through the hazard ratio (HR) and odds ratio (OR) estimated by the Cox proportional hazard regression model and logistic model, respectively. A bivariate Plackett copula model[15] was used to estimate the HR for PFS and the OR for complete response end points when patient-level correlation was estimated simultaneously. Analyses were performed with SAS 9.4 (SAS Institute, Cary, NC) and R version 2.14.0 (The R Project) software.

Surrogacy evaluation. At the trial level, two commonly used surrogacy measures were considered: R^2_{Copula}[15] and R^2_{WLS}[16,17] (which is based on the weighted least square [WLS] regression method), where R^2_{Copula} takes into account patient-level correlation between the two end points and R^2_{WLS} does not. The predefined rule for declaring trial-level surrogacy required R^2_{WLS} or $R^2_{Copula} \geq 0.80$ with a lower 95% CI bound > 0.6 and neither estimate < 0.7. The threshold of 0.8 was defined as a conservative approach in this setting, wherein precedent is limited.[16-18] Supplemental trial-level surrogacy measures included the surrogate threshold effect[19]; the minimum treatment effect on the surrogate required to confidently predict a significant treatment effect on PFS in a future trial; and concordance of significance,[20] which was defined as the proportion of trials with the same conclusions for both the surrogate and the true end point (ie, both significant or both nonsignificant for treatment effect). At the patient level, the global OR for comparing PFS status beyond a particular time point between patients with and without complete response was estimated through the bivariate Plackett copula model[15]; patient-level correlation is a supportive but not a sufficient condition for surrogacy validation.

Sensitivity and Subpopulation Analysis

Leave-one-out cross-validation, which compares the predicted with the observed log(HR)s on PFS on the basis of the estimated trial-level model that leaves one trial out at a time, was used to assess the prediction performance of the regression model. Leave-one-out estimation, which re-estimated the R^2 when one trial was excluded at a time, identified potential influential trials. Secondary calculation rules (Appendix Fig A1, online only) to account for the typical course of follicular lymphoma in patients who underwent first-line treatment provided a sensitivity analysis. Further

sensitivity analyses were both unconfirmed complete response and complete response in the calculations. Surrogacy was further examined within subpopulations defined by treatment type (ie, studies with and without rituximab), study design (studies that randomly assigned patients before induction versus before maintenance treatment), and key patient characteristics (high-risk versus low/intermediate-risk Follicular Lymphoma International Prognostic Index [FLIPI] score).

RESULTS

Patient Characteristics

Three hundred forty-six references were individually examined; 29 studies met the required selection criteria (Fig 1). The owners of these 29 studies were contacted, and 22 provided individual patient data. Seven were subsequently excluded for lack of sufficient data to derive 30-month complete response. Two of 15 studies had > 20% missing complete response status after clinical calculation rules were applied and were subsequently excluded from further analyses. Therefore, 13 studies (eight induction, five maintenance) in 3,837 patients were analyzed, with all exclusions solely based on data availability irrespective of the actual trial 30-month complete response and PFS results. Table 1 lists the trial-level characteristics of the 13 studies.[3,9,21-33] Median patient age was 56 years (range, 19 to 90 years); 50.3% were male, 57.6% had Eastern Cooperative Oncology Group performance status of 0 at enrollment, 44.0% had a high FLIPI score, and 67.8% had Ann Arbor stage IV disease. Overall, the excluded studies were conducted during the same era as the those included, with similar patient characteristics. Patient treatment assignment was analyzed according to the intention-to-treat principle. All examined covariates were well balanced between the experimental and the control arms (Table 2). Median follow-up time among patients without progressive disease ranged from 30 to 151 months between studies.

Surrogacy Analysis for Principal Surrogate End Point Candidate (30-Month Complete Response)

Trial-level correlation between CR30 and PFS. Figure 2A shows the trial-level association between treatment effects on CR30 and PFS. A strong correlation was observed between the treatment effects on the two end points at the trial level (R^2_{WLS}, 0.88 [95% CI, 0.77 to 0.96]; R^2_{Copula}, 0.86 [95% CI, 0.72 to 1.00]), which met the predefined criteria for surrogacy. Twelve (93.2%) of the 13 trials had concordance of significance. The estimated surrogate threshold effect was 1.56, which indicated that an observed OR

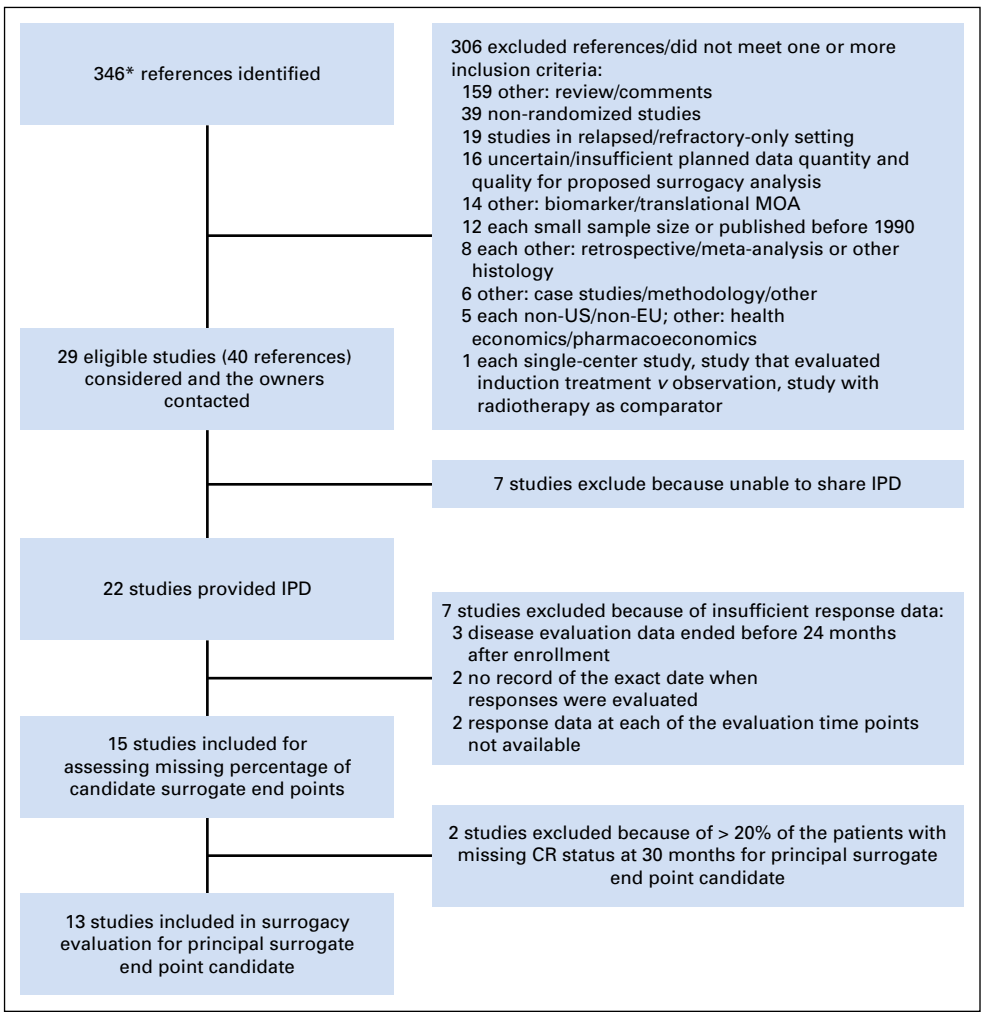

Fig 1. Flowchart of study selection. *Three hundred forty-two references were identified through the initial September 2011 Medline database search, and four references were published after the initial literature search. CR, complete response; EU, European Union; IPD, individual patient data; MOA, mechanism of action.

Table 1. Phase III Randomized Studies in Follicular Lymphoma That Met Selection Criteria and Provided Sufficient Data for Inclusion in the Analysis

Research Group*	First Author	Study Identifier†	Start of Accrual	Arm	No.‡	Induction	Maintenance	Random Assignment at Induction or Maintenance?	Trial Included Rituximab?	Median Follow-up Time to Progression (Q1, Q3), months§
CALGB	Peterson[21]	CALGB 7951	1980	Control	86	CHOP-B	CHOP-B	Induction	No	151 (102-189)
				Exp	103	Cyclophosphamide	Cyclophosphamide			
EORTC	Hagenbeek[22]	EORTC 20921, ERP_165	1993	Control	117	CVP	—	Induction	No	55 (47-72)
				Exp	114	Fludarabine	—			
GOELAMS	Gyan[23] Deconinck[24]	GOELAMS 064, NCT00696735	1994	Control	81	CHVP-IFN-α	CHVP-IFN-α	Induction	No	82 (65-106)
				Exp	85	VCAP/ASCT	Observation			
GLSG	Nickenig[25]	GLSG1996	1996	Control	146	MCP	ASCT/IFN-α	Induction	No	108 (49-137)
				Exp	362	CHOP	ASCT/IFN-α			
ECOG	Hochster[9]	ECOG E1496, NCT00003204	1998	Control	113	CVP	Observation	Maintenance	Yes	98 (83-116)
				Exp	115	CVP	Rituximab			
OSHO	Herold[26]	OSHO-39, M39023, NCT00269113	1998	Control	96	MCP	IFN-α	Induction	Yes	51 (41-54)
				Exp	105	R-MCP	IFN-α			
SAKK‖	Ghielmini[27] Martinelli[28]	SAKK 35/98, NCT00003280	1998	Control	23	Rituximab	Observation	Maintenance	Yes	105 (89-118)
				Exp	22	Rituximab	Rituximab			
GLSG	Hiddemann[29]	GLSG2000	2000	Control	290	CHOP	ASCT/IFN-α	Induction	Yes	84 (62-100)
				Exp	292	R-CHOP	ASCT/IFN-α			
Roche	Marcus[30]	M39021	2000	Control	160	CVP	—	Induction	Yes	95 (79-101)
				Exp	162	R-CVP	—			
Favrille/ MMRGlobal	Freedman[31]	FavId06, NCT00089115	2004	Control	130	Rituximab	Placebo	Maintenance	Yes	30 (24-35)
				Exp	127	Rituximab	Idiotype vaccine			
FIL	Vitolo[32]	ML17638, NCT01144364	2004	Control	101	R-FND	Observation	Maintenance	Yes	43 (41-43)
				Exp	101	R-FND	Rituximab			
GELA	Salles[3]	PRIMA/MO18264, NCT00140582	2004	Control	513	R-CHOP/R-CVP/R-FCM	Observation	Maintenance	Yes	54 (48-59)
				Exp	505	R-CHOP/R-CVP/R-FCM	Rituximab			
NLG	Kimby[33]	ML16865, NCT01609010	2004	Control	117	Rituximab	Rituximab	Induction	Yes	59 (44-79)
				Exp	111	R + IFN-α	R + IFN-α			

Abbreviations: ASCT, autologous stem cell transplantation; CALGB, Cancer and Leukemia Group B; CHOP, cyclophosphamide, doxorubicin, vincristine, prednisone; CHOP-B, cyclophosphamide, doxorubicin, vincristine, prednisone, bleomycin; CHVP, cyclophosphamide, doxorubicin, etoposide/teniposide, prednisone/prednisolone; CVP, cyclophosphamide, vincristine, prednisone; ECOG, Eastern Cooperative Oncology Group; EORTC, European Organization for Research and Treatment of Cancer; Exp, experimental; FavId06, Phase III Trial of Favid and Gmcsf [granulocyte-macrophage colony-stimulating factor] vs Placebo and Gmcsf Following Rituximab; FCM, fludarabine, cyclophosphamide, mitoxantrone; FIL, Fondazione Italiana Linfomi; FND, fludarabine, mitoxantrone, dexamethasone; GELA, Study Group of the Adult Lymphoma; GLSG, German Low-Grade Lymphoma Study Group; GOELAMS, French Acute Leukemia and Blood Diseases West-East Group; IFN-α: interferon alfa; MCP, mitoxantrone, chlorambucil, prednisone; NLG, Nordic Lymphoma Group; OSHO, Ostdeutsche Studiengruppe Hamato-Onkologie; PRIMA, Advanced Ovarian Cancer Following Response on Front-Line Platinum-Based Chemotherapy; R, rituximab; SAKK, Swiss Group for Clinical Cancer Research; VCAP, vincristine, cyclophosphamide, doxorubicin, prednisone.

*OSHO-19 and FL2000 were further excluded because > 20% of the patients had missing complete response rate at 30 months data per primary calculation rules.

†ClinicalTrials.gov identifier provided where available.

‡Number of randomly assigned untreated patients with follicular lymphoma on the basis of the data transferred.

§Calculated among patients with progression-free disease.

‖On the basis of the SAKK 35/98 publication: 151 patients (78 in observation arm; 73 in rituximab arms) with a complete response, a partial response, or stable disease were randomly assigned of whom 51 (25 observation; 26 rituximab) were previously untreated, hence meeting the initial study inclusion criteria of ≥ 50 patients with follicular lymphoma. In the data transferred to the statistics and data center, only 45 randomly assigned untreated patients with follicular lymphoma (23 observation; 22 rituximab) were identified.

Table 2. Patient Characteristics on the Basis of 13 Trials Included in the Primary Analysis

Characteristic	Control, No. (%)	Experimental, No. (%)	Total, No. (%)	*P*
No. of patients	1,988	1,849	3,837	
Age (continuous)				.0306*
Mean (SD)	54.92 (11.23)	55.71 (11.33)	55.30 (11.29)	
Median	55.67	56.51	56.00	
Range	21.00-85.67	18.96-90.10	18.96-90.10	
Age (categorical), years				.0268†
< 50	620 (31.2)	552 (29.9)	1,172 (30.5)	
50-59	668 (33.6)	561 (30.3)	1,229 (32.0)	
60-69	507 (25.5)	537 (29.0)	1,044 (27.2)	
≥ 70	193 (9.7)	199 (10.8)	392 (10.2)	
Sex				.5590†
Female	979 (49.2)	928 (50.2)	1,907 (49.7)	
Male	1,009 (50.8)	921 (49.8)	1,930 (50.3)	
ECOG performance status‡				.0972†
Missing	84	115	199	
0	1,065 (55.9)	1,030 (59.4)	2,095 (57.6)	
1	752 (39.5)	626 (36.1)	1,378 (37.9)	
2+	87 (4.6)	78 (4.5)	165 (4.5)	
FLIPI category§				.9323†
Missing	229	250	479	
Low (0-1)	311 (17.7)	288 (18.0)	599 (17.8)	
Intermediate (2)	677 (38.5)	606 (37.9)	1,283 (38.2)	
High (3-5)	771 (43.8)	705 (44.1)	1,476 (44.0)	
Ann Arbor stage‖				.5672†
Missing	1	4	5	
I/II	106 (5.3)	93 (5.0)	199 (5.2)	
III	522 (26.3)	512 (27.8)	1,034 (27.0)	
IV	1,359 (68.4)	1,240 (67.2)	2,599 (67.8)	

Abbreviations: ECOG, Eastern Cooperative Oncology Group; FLIPI, Follicular Lymphoma International Prognostic Index; SD, standard deviation.
*Unequal variance *t* test.
†χ^2 test.
‡ECOG performance status describes a patient's level of functioning in terms of ability to care for self, daily activity, and physical ability on a scale of 0 (fully active) to 5 (dead).
§FLIPI is a validated index that is based on five clinical factors (age > 60 years, disease stage III or IV, more than four lymph node groups involved, serum hemoglobin < 12 g/dL, and serum lactate dehydrogenase greater than the upper limit of normal).[39] If present, each factor is assigned 1 point, and the total FLIPI score divides patients into three levels of risk that are prognostic for overall survival: 0 to 1 points (low risk), 2 points (intermediate risk), and 3 to 5 points (high risk).
‖The Ann Arbor staging system for lymphomas is determined by the location and extent of the disease and ranges from stage I (cancer is located in a single region) to stage IV (diffuse or disseminated involvement).

≥ 1.56 for CR30 would predict a significant treatment effect on PFS in a future trial, which translates to an absolute improvement in the CR30 of 11%, assuming a control arm CR30 of 50%.

Patient-level correlation between CR30 and PFS. The global OR was 11.8 (95% CI, 10.0 to 13.7), which indicates substantially higher odds of remaining alive and progression free beyond a particular time point for patients who achieve complete response at 30 months compared with those who do not.

Sensitivity analyses. Leave-one-out cross-validation demonstrated consistency between observed and predicted PFS treatment effects for each trial on the basis of the CR30 (Fig 2B). To identify potentially highly influential trials, one trial at a time was excluded to re-estimate the R^2 measures (Appendix Fig A2, online only); none of the R^2 estimates were < 0.8, and none of the lower bounds of the 95% CI were < 0.6. By applying secondary clinical calculation rules, the surrogacy of CR30 was consistent (R^2_{WLS}, 0.88 [95% CI, 0.79 to 0.97]; R^2_{Copula}, 0.86 [95% CI, 0.71 to 1.00]). Eight studies recorded unconfirmed complete response. When including both complete response and unconfirmed complete response, slightly stronger trial-level surrogacy was observed for CR30 (R^2_{WLS}, 0.96 [95% CI, 0.90 to 1.00]; R^2_{Copula}, 0.96 [95% CI, 0.91 to 1.00]).

Subgroup analyses. Subpopulation analyses that separately examined trials that contained or did not contain rituximab, induction trials, maintenance trials, and patients with high FLIPI scores showed consistently high levels of surrogacy, with R^2 estimates from 0.8 to 0.9. However, although patient-level correlation remained strong, among the subset of patients with low-to-intermediate FLIPI scores, the R^2 statistics were moderate (range, 0.5 to 0.6; Table 3).

Surrogacy Evaluation for Secondary Surrogate End Point Candidate (CR24)

Because CR30 met the predefined criteria, the surrogacy of CR24 was evaluated. Four studies had > 20% missing data for 24-month complete response and were excluded. Among the remaining 11 studies (2,728 patients), patient-level correlation remained high (global OR, 8.27; 95% CI, 6.82 to 9.71). However, the strong correlation between treatment effects on CR24 and PFS was demonstrated only by R^2_{WLS} (0.84; 95% CI, 0.63 to 0.95), not R^2_{Copula} (0.67; 95% CI, 0.35 to 0.99). In analyses that excluded one small outlier study (Fig 2C), both R^2_{WLS} (0.86; 95% CI, 0.70 to 0.97) and R^2_{Copula} (0.83; 95% CI, 0.65 to 1.00) demonstrated strong

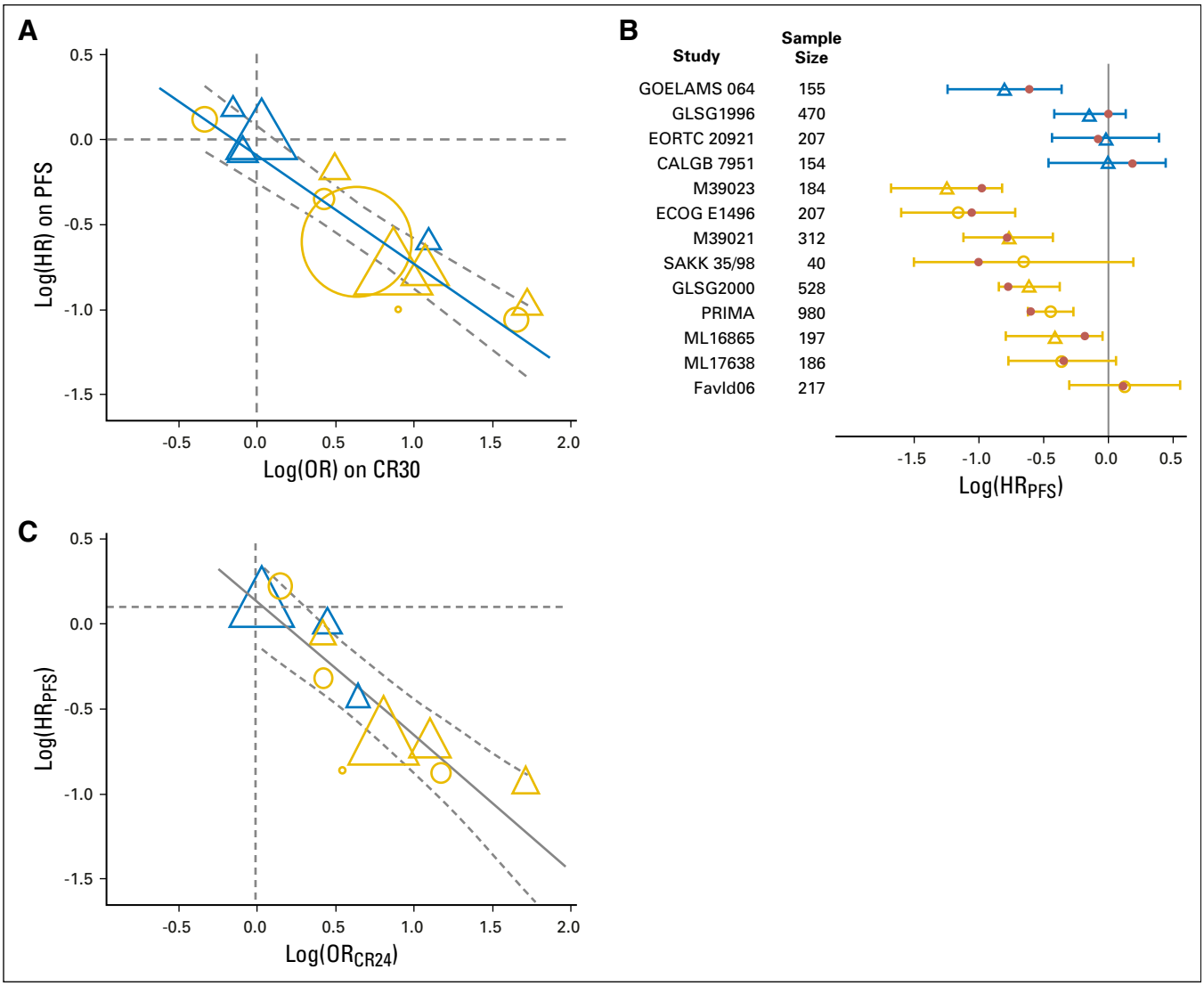

Fig 2. (A) Trial-level correlation between complete response rate at 30 months (CR30) after initiation of induction treatment (per primary calculation rules) and progression-free survival (PFS). Thirteen studies are included. Gold indicates rituximab trials; blue, nonrituximab trials; triangle, induction trial; and circle, maintenance trial. The size of the circles and triangles is proportional to the sample size. The fitted weighted least squares regression line (solid line), is log(hazard ratio $[HR]_{PFS}$) = −0.093 – 0.636 × log(odds ratio $[OR]_{CR30}$). The dashed lines indicate 95% prediction limits. The horizontal dashed line corresponds to the $log(HR_{PFS})$ of 0 (ie, HR of 1). The vertical dashed line corresponds to the $log(OR_{CR30})$ of 0 (ie, OR of 1). (B) Assessment of the prediction of $log(HR_{PFS})$ on the basis of the estimated regression model at the trial level by leave-one-out cross-validation (CR30 based on primary calculation rules). Gold indicates rituximab trials; blue, nonrituximab trials; triangle, induction trial; circle, maintenance trial. For each trial, the open circles and triangles are the predicted $log(HR_{PFS})$ based on the estimated weighted least squares regression line at the trial level after removing the trial listed; the horizontal bars are the 95% prediction intervals. The solid red circles are the observed $log(HR_{PFS})$ of the trial. (C) Trial-level correlation between complete response rate at 24 months (CR24) after initiation of induction treatment and PFS (per primary calculation rules). Eleven studies are included. Gold indicates rituximab trials; blue, nonrituximab trials; triangle, induction trial; and circle, maintenance trial. The size of the circles and triangles is proportional to the sample size. The fitted weighted least squares regression line (solid line), is $log(HR_{PFS})$ = 0.043 – 0.726 × log (OR_{CR24}). The dashed lines indicate 95% prediction limits. The horizontal dashed line corresponds to the $log(HR_{PFS})$ of 0 (ie, HR of 1). The vertical dashed line corresponds to the $log(OR_{CR24})$ of 0 (ie, OR of 1).

trial-level correlation. Because this exclusion was post hoc, the results for CR24 are considered hypothesis generating only.

DISCUSSION

This pooled analysis assembled individual patient data from 13 randomized trials in follicular lymphoma initiated worldwide from 1980 to 2007. To our knowledge, this analysis is the first to be based on integrated individual patient data from randomized controlled trials in lymphoma. Unlike literature-based meta-analysis, individual patient data ensure the consistent calculation of end points and hence, the consistent interpretation of the within-trial treatment effects across all studies. The analysis demonstrated that treatment effects on CR30 strongly predict treatment effects on PFS. The results are highly consistent across various surrogacy estimation methods and sensitivity analyses. The strong association was maintained irrespective of the inclusion of rituximab in the

Table 3. Surrogacy Measures of Principal Surrogate Candidate: Complete Response Rate at 30 Months After Initiation of Induction Treatment on the Basis of Primary Calculation Rules

Analysis Population	No. of Trials (No. of Patients)	Trial-Level Surrogacy		Patient-Level Surrogacy
		R^2_{WLS} (95% CI)	R^2_{Copula} (95% CI)	Global OR* (95% CI)
Overall	13 (3,837)	0.88 (0.77 to 0.96)	0.86 (0.72 to 1.00)	11.84 (10.03 to 13.65)
Within subpopulations defined by trials with or without rituximab				
Rituximab trial†	9 (2,851)	0.85 (0.62 to 0.97)	0.80 (0.56 to 1.00)	11.08 (9.13 to 13.03)
Nonrituximab trial	4 (986)	0.91 (0.05 to 1.00)	0.96 (0.90 to 1.00)	14.40 (9.96 to 18.84)
Within subpopulations defined by study type				
Induction trial‡	8 (2,207)	0.89 (0.75 to 0.98)	0.89 (0.74 to 1.00)	10.34 (8.27 to 12.41)
Maintenance trial‡	5 (1,630)	0.93 (0.84 to 1.00)	0.89 (0.71 to 1.00)	14.14 (10.82 to 17.46)
Within subpopulation defined by FLIPI§				
High‖	9 (1,415)	0.87 (0.68 to 0.98)	0.73 (0.42 to 1.00)	14.10 (10.43 to 17.78)
Low to intermediate	10 (1,882)	0.45 (0.02 to 0.93)	0.57 (0.17 to 0.97)	9.00 (7.07 to 10.92)

Abbreviations: Copula, bivariate Plackett copula model; FLIPI, Follicular Lymphoma International Prognostic Index; HR, hazard ratio; OR, odds ratio; WLS, weight least squares.

*The global OR for progression-free survival status beyond a particular time point that compares responders versus nonresponders was estimated through the bivariate Plackett copula model. The higher the value, the stronger the association. A lower bound of the 95% CI > 1 indicates a significant association.

†Rituximab trial refers to studies with at least one arm with rituximab.

‡Induction (maintenance) trial refers to studies with random assignment before start of induction (maintenance) treatment.

§FLIPI is a validated index that is based on five clinical factors (age > 60 years, disease stage III or IV, more than four lymph node groups involved, serum hemoglobin < 12 g/dL, and serum lactate dehydrogenase greater than the upper limit of normal).[39] If present, each factor is assigned 1 point, and the total FLIPI score divides patients into three levels of risk that are predictive of overall survival: 0 to 1 points (low risk), 2 points (intermediate risk), and 3 to 5 points (high risk). Cancer and Leukemia Group B (CALGB) 7951, ERP_165, and Swiss Group for Clinical Cancer Research (SAKK) 35/98 were not included because FLIPI scores were not supplied.

‖Further excluding the FavId06 study because among patients with high FLIPI scores, all were in the experimental arm and classified as having noncomplete response at 30 months.

regimen or whether the trials involved random assignment at induction or maintenance. When including both complete response and unconfirmed complete response (available in eight studies), a slightly stronger trial-level surrogacy was observed.

Seven of the 13 included studies began accrual after 1999, and all used the 1999 National Cancer Institute–sponsored International Working Group or similar response criteria.[34] All studies required rigorous response assessment schedules, with clinical and physical examination and CT imaging. Median duration of complete response was 41.6 months (range, 2.3 to 175.2 months), with 96.2%, 83.8%, and 62.2% of complete responses lasting > 1, 2, and 3 years, respectively. The end point of CR30 thus represents a durable treatment-induced complete response, which supports its potential use as a primary end point in future trials in which the treatment goal is the achievement and maintenance of complete response. The relationship between CR30 and PFS was maintained across all evaluated trials, including induction-only trials, which suggests that even if standard-of-care treatment changes from a 30-month total course, the end point will be robust for therapeutic strategies that seek durable complete response. These findings support CR30 as appropriate for use as a primary study end point in patients with previously untreated follicular lymphoma, with the intent of bringing novel therapies to this patient population years before PFS results are available.

End-of-treatment response assessment may be more accurate with positron emission tomography (PET)/CT imaging than with CT imaging alone. Retrospective[35] and prospective[36] studies have demonstrated improved PFS prognostication in patients with PET-negative relative to PET-positive findings at the end of treatment. PET/CT imaging has been recommended for use in the evaluation of response in PET-avid lymphoma, including follicular lymphoma,[37] whereby partial response and unconfirmed complete response according to International Working Group 1999 criteria[34] are classified as complete metabolic response if residual lesions are PET negative. PET negativity at the end of treatment requires further investigation as a potential surrogate end point, but because these criteria are a modification rather than a major redefinition, we expect that the relationship between CR30 and PFS would remain with PET evaluation.

Only one of the two primary trial-level surrogacy measures for CR24 met the predefined qualification criteria. A post hoc analysis that excluded an outlier study resulted in both trial-level surrogacy measures showing strong correlation, consistent with a recent cohort study of patients with follicular lymphoma treated with chemoimmunotherapy.[11] Additional data are required to reconsider surrogacy of the complete response rate at this assessment time point. Because CR24 did not meet our prespecified criteria, no additional end points were considered as a part of our formal end point validation. Subsequent hypothesis-generating analyses of additional possible end points derived by using this rich data set are ongoing. The extended overall survival of patients, which is the reason that PFS is the current standard end point, precludes a meaningful assessment of the relationship between CR30 and overall survival.

Patients' clinical and disease characteristics affected the estimated surrogacy of CR30. A strong correlation existed between CR30 and PFS among patient groups with high FLIPI scores, whereas the correlation among patients with low to intermediate FLIPI scores was less strong. The reduced correlation is likely due to the prognostic heterogeneity in the low to intermediate FLIPI groups. Similar findings were observed when CR30 surrogacy was evaluated on the basis of patients with stage IV disease (n = 2,585; R^2_{WLS}, 0.92 [95% CI, 0.85 to 0.97]; R^2_{Copula}, 0.94 [95% CI, 0.87 to 1.00]) versus stage I to III disease (n = 1,207; R^2_{WLS}, 0.58 [95% CI, 0.25 to 0.87]; R^2_{Copula}, 0.59 [95% CI, 0.24 to 0.95]). In ongoing phase III trials, the majority of patients have stage III or IV disease

because the treatment approach is similar for both stages. To inform the use of CR30 as a surrogate end point in future trials, we estimated trial-level correlation in patients with stage III and IV disease and identified high correlation (R^2_{WLS}, 0.87 [95% CI, 0.75 to 0.97]; R^2_{Copula}, 0.91 [95% CI, 0.81 to 1.00]). Because we expect patients with high risk or high tumor burden to be the main study population for further first-line follicular lymphoma treatment trials in which the achievement of durable complete response is relevant, the consistently strong correlation between CR30 and PFS observed in the high-risk FLIPI and stage III to IV populations supports CR30 as a surrogate end point in such trials. Furthermore, the median age of patients included in this analysis is younger than that of patients with follicular lymphoma in the general population. This anticipated phenomenon is associated with clinical trial enrollment and does not preclude the use of CR30 as a surrogate end point for follicular lymphoma trials, but it limits the extrapolation of CR30 as a predictor of PFS for patients treated in general clinical practice.

As with any surrogate end point, the use of CR30 in future trials must be comprehensively considered in the context of disease population and treatment type. The findings are based on trials of chemotherapy with or without rituximab and on treatments that lasted $\leq$ 30 months. We recommend CR30 as an appropriate primary end point for trials that evaluate novel therapies that aim to maintain and/or increase durable complete response in patients with previously untreated follicular lymphoma. Future updates with data from newly maturing trials will be important to ensure continued applicability of these findings, and trials should continue to collect robust PFS and overall survival data to fully understand long-term treatment effects. Other surrogate end point candidates, such as complete response or PFS status at earlier time points are to be explored in future studies within the FLASH collaboration. These findings also demonstrate the value and potential of international data sharing consortia to address critical issues that one trial or group cannot address alone.[17,38]

In conclusion, this pooled analysis of randomized chemotherapy, immunotherapy, or chemoimmunotherapy trials demonstrates that complete response at 30 months after initiation of induction treatment may serve as a surrogate end point for PFS in first-line follicular lymphoma treatment trials.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Disclosures provided by the authors are available with this article at ascopubs.org/journal/jco.

AUTHOR CONTRIBUTIONS

Conception and design: Qian Shi, Christopher R. Flowers, Robert Marcus, Umberto Vitolo, Tina Nielsen, Nancy Valente, Nathan H. Fowler, Gilles Salles, Daniel J. Sargent
Provision of study materials or patients: Michele Ghielmini, Gilles Salles
Collection and assembly of data: Qian Shi, Christopher R. Flowers, Wolfgang Hiddemann, Michael Herold, Anton Hagenbeek, Eva Kimby, Howard Hochster, Umberto Vitolo, Bruce A. Peterson, Emmanuel Gyan, Michele Ghielmini, Sabine De Bedout, Nancy Valente, Nathan H. Fowler, Eva Hoster, Marco Ladetto, Franck Morschhauser, Emanuele Zucca, Gilles Salles, Daniel J. Sargent
Data analysis and interpretation: All authors
Manuscript writing: All authors
Final approval of manuscript: All authors
Accountable for all aspects of the work: All authors

REFERENCES

1. Howlader N, Noone A, Krapcho M, et al: SEER Cancer Statistics Review, 1975-2012. National Cancer Institute, Bethesda, MD, 2015.
2. Nastoupil LJ, Sinha R, Byrtek M, et al: Comparison of the effectiveness of frontline chemoimmunotherapy regimens for follicular lymphoma used in the United States. Leuk Lymphoma 56: 1295-1302, 2015
3. Salles G, Seymour JF, Offner F, et al: Rituximab maintenance for 2 years in patients with high tumour burden follicular lymphoma responding to rituximab plus chemotherapy (PRIMA): A phase 3, randomised controlled trial. Lancet 377:42-51, 2011
4. Katz R: Biomarkers and surrogate markers: An FDA perspective. NeuroRx 1:189-195, 2004
5. Bachy E, Brice P, Delarue R, et al: Long-term follow-up of patients with newly diagnosed follicular lymphoma in the prerituximab era: Effect of response quality on survival—A study from the Groupe d'Etude des Lymphomes de l'Adulte. J Clin Oncol 28: 822-829, 2010
6. Buckstein R, Pennell N, Berinstein NL: What is remission in follicular lymphoma and what is its relevance? Best Pract Res Clin Haematol 18:27-56, 2005
7. Johnson PW, Rohatiner AZ, Whelan JS, et al: Patterns of survival in patients with recurrent follicular lymphoma: A 20-year study from a single center. J Clin Oncol 13:140-147, 1995
8. López-Guillermo A, Montserrat E, Bosch F, et al: Applicability of the International Index for Aggressive Lymphomas to patients with low-grade lymphoma. J Clin Oncol 12:1343-1348, 1994
9. Hochster H, Weller E, Gascoyne RD, et al: Maintenance rituximab after cyclophosphamide, vincristine, and prednisone prolongs progression-free survival in advanced indolent lymphoma: Results of the randomized phase III ECOG1496 Study. J Clin Oncol 27:1607-1614, 2009
10. Rummel MJ, Niederle N, Maschmeyer G, et al: Bendamustine plus rituximab is superior in respect of progression free survival and CR rate when compared to CHOP plus rituximab as first-line treatment of patients with advanced follicular, indolent, and mantle cell lymphomas: Final results of a randomized phase III study of the StiL (Study Group Indolent Lymphomas, Germany). Blood 114, 2009 (abstr 405)
11. Casulo C, Byrtek M, Dawson KL, et al: Early relapse of follicular lymphoma after rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone defines patients at high risk for death: An analysis from the National LymphoCare Study. J Clin Oncol 33:2516-2522, 2015
12. Lee L, Wang L, Crump M: Identification of potential surrogate end points in randomized clinical trials of aggressive and indolent non-Hodgkin's lymphoma: Correlation of complete response, time-to-event and overall survival end points. Ann Oncol 22: 1392-1403, 2011
13. Saville MW, Leonard JP, Hainsworth JD, et al: Role of different frontline regimens in achieving complete response in follicular lymphoma: A meta-analysis of CR rate and its relation to hazard rate for disease progression. Blood 108, 2006 (abstr 2754)
14. Salles G, Mounier N, de Guibert S, et al: Rituximab combined with chemotherapy and interferon in follicular lymphoma patients: Results of the GELA-GOELAMS FL2000 study. Blood 112:4824-4831, 2008
15. Burzykowski T, Molenberghs G, Buyse M: The validation of surrogate end points by using data from randomized clinical trials: A case-study in advanced colorectal cancer. J R Stat Soc Ser A Stat Soc 167: 103-124, 2004
16. Sargent DJ, Patiyil S, Yothers G, et al: End points for colon cancer adjuvant trials: Observations and recommendations based on individual patient data from 20,898 patients enrolled onto 18 randomized trials from the ACCENT Group. J Clin Oncol 25:4569-4574, 2007
17. Sargent DJ, Wieand HS, Haller DG, et al: Disease-free survival versus overall survival as a primary end point for adjuvant colon cancer studies: Individual patient data from 20,898 patients on 18 randomized trials. J Clin Oncol 23:8664-8670, 2005

18. Burzykowski T, Molenberghs G, Buyse M, et al: Validation of surrogate end points in multiple randomized clinical trials with failure time end points. Appl Stat 50:405-422, 2001

19. Burzykowski T, Buyse M: Surrogate threshold effect: An alternative measure for meta-analytic surrogate endpoint validation. Pharm Stat 5:173-186, 2006

20. Begg CB, Leung DHY: On the use of surrogate end points in randomized trials. J R Stat Soc Ser A Stat Soc 163:15-28, 2000

21. Peterson BA, Petroni GR, Frizzera G, et al: Prolonged single-agent versus combination chemotherapy in indolent follicular lymphomas: A study of the Cancer and Leukemia Group B. J Clin Oncol 21: 5-15, 2003

22. Hagenbeek A, Eghbali H, Monfardini S, et al: Phase III intergroup study of fludarabine phosphate compared with cyclophosphamide, vincristine, and prednisone chemotherapy in newly diagnosed patients with stage III and IV low-grade malignant non-Hodgkin's lymphoma. J Clin Oncol 24:1590-1596, 2006

23. Gyan E, Foussard C, Bertrand P, et al: High-dose therapy followed by autologous purged stem cell transplantation and doxorubicin-based chemotherapy in patients with advanced follicular lymphoma: A randomized multicenter study by the GOELAMS with final results after a median follow-up of 9 years. Blood 113:995-1001, 2009

24. Deconinck E, Foussard C, Milpied N, et al: High-dose therapy followed by autologous purged stem-cell transplantation and doxorubicin-based chemotherapy in patients with advanced follicular lymphoma: A randomized multicenter study by GOELAMS. Blood 105:3817-3823, 2005

25. Nickenig C, Dreyling M, Hoster E, et al: Combined cyclophosphamide, vincristine, doxorubicin, and prednisone (CHOP) improves response rates but not survival and has lower hematologic toxicity compared with combined mitoxantrone, chlorambucil, and prednisone (MCP) in follicular and mantle cell lymphomas: Results of a prospective randomized trial of the German Low-Grade Lymphoma Study Group. Cancer 107:1014-1022, 2006

26. Herold M, Haas A, Srock S, et al: Rituximab added to first-line mitoxantrone, chlorambucil, and prednisolone chemotherapy followed by interferon maintenance prolongs survival in patients with advanced follicular lymphoma: An East German Study Group Hematology and Oncology Study. J Clin Oncol 25:1986-1992, 2007

27. Ghielmini M, Schmitz SF, Cogliatti SB, et al: Prolonged treatment with rituximab in patients with follicular lymphoma significantly increases event-free survival and response duration compared with the standard weekly x 4 schedule. Blood 103:4416-4423, 2004

28. Martinelli G, Schmitz SF, Utiger U, et al: Long-term follow-up of patients with follicular lymphoma receiving single-agent rituximab at two different schedules in trial SAKK 35/98. J Clin Oncol 28: 4480-4484, 2010

29. 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 106:3725-3732, 2005

30. Marcus R, Imrie K, Solal-Celigny P, et al: Phase III study of R-CVP compared with cyclophosphamide, vincristine, and prednisone alone in patients with previously untreated advanced follicular lymphoma. J Clin Oncol 26:4579-4586, 2008

31. Freedman A, Neelapu SS, Nichols C, et al: Placebo-controlled phase III trial of patient-specific immunotherapy with mitumprotimut-T and granulocyte-macrophage colony-stimulating factor after rituximab in patients with follicular lymphoma. J Clin Oncol 27: 3036-3043, 2009

32. Vitolo U, Ladetto M, Boccomini C, et al: Rituximab maintenance compared with observation after brief first-line R-FND chemoimmunotherapy with rituximab consolidation in patients age older than 60 years with advanced follicular lymphoma: A phase III randomized study by the Fondazione Italiana Linfomi. J Clin Oncol 31:3351-3359, 2013

33. Kimby E, Östenstad B, Brown P, et al: Two courses of four weekly infusions of rituximab with or without interferon-α2a: Final results from a randomized phase III study in symptomatic indolent B-cell lymphomas. Leuk Lymphoma 56:2598-2607, 2015

34. Cheson BD, Horning SJ, Coiffier B, et al: Report of an international workshop to standardize response criteria for non-Hodgkin's lymphomas. J Clin Oncol 17:1244, 1999 [Erratum: J Clin Oncol 18, 2351, 2000]

35. Trotman J, Fournier M, Lamy T, et al: Positron emission tomography-computed tomography (PET-CT) after induction therapy is highly predictive of patient outcome in follicular lymphoma: Analysis of PET-CT in a subset of PRIMA trial participants. J Clin Oncol 29:3194-3200, 2011

36. Dupuis J, Berriolo-Riedinger A, Julian A, et al: Impact of [(18)F]fluorodeoxyglucose positron emission tomography response evaluation in patients with high-tumor burden follicular lymphoma treated with immunochemotherapy: A prospective study from the Groupe d'Etudes des Lymphomes de l'Adulte and GOELAMS. J Clin Oncol 30:4317-4322, 2012

37. 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 32:3059-3068, 2014

38. Shi Q, de Gramont A, Grothey A, et al: Individual patient data analysis of progression-free survival versus overall survival as a first-line end point for metastatic colorectal cancer in modern randomized trials: Findings from the Analysis and Research in Cancers of the Digestive System database. J Clin Oncol 33:22-28, 2015

39. Solal-Céligny P, Pascal R, Colombat P, et al: Follicular lymphoma international prognostic index. Blood 104:1258-1265, 2004

Affiliations

Qian Shi and **Daniel J. Sargent**, Mayo Clinic, Rochester; **Bruce A. Peterson**, University of Minnesota, Minneapolis, MN; **Christopher R. Flowers**, Winship Cancer Institute of Emory University, Atlanta, GA; **Wolfgang Hiddemann** and **Eva Hoster**, Ludwig-Maximilians University Hospital, Munich; **Michael Herold**, HELIOS Kliniken, Erfurt, Germany; **Robert Marcus**, Addenbrooke's Hospital, Cambridge, United Kingdom; **Anton Hagenbeek**, Academic Medical Center, Amsterdam, the Netherlands; **Eva Kimby**, Karolinska Institutet, Stockholm, Sweden; **Howard Hochster**, Yale Cancer Center, New Haven, CT; **Umberto Vitolo**, Azienda Ospedaliero-Universitaria Città della Salute e della Scienza di Torino, Turin; **Marco Ladetto**, Azienda Ospedaliera SS Antonio e Biagio e Cesare Arrigo, Alessandria, Italy; **Emmanuel Gyan**, University Hospital, Tours; **Franck Morschhauser**, Service Université de Lille 2, Lille; **Gilles Salles**, Université Claude Bernard, Pierre Bénite, France; **Michele Ghielmini** and **Emanuele Zucca**, Oncology Institute of Southern Switzerland, Bellinzona; **Michele Ghielmini** and **Emanuele Zucca**, Swiss Group for Clinical Cancer Research, Bern; **Tina Nielsen**, F. Hoffmann-La Roche, Basel; **Sabine De Bedout**, Celgene, Boudry, Switzerland; **Tommy Fu**, Celgene, Summit, NJ; **Nancy Valente**, Genentech, South San Francisco, CA; and **Nathan H. Fowler**, University of Texas MD Anderson Cancer Center, Houston, TX.

Support

Supported by grants from Celgene and Roche. Data were transmitted directly from the original study cooperative groups to the Mayo Clinic. Celgene and Roche further jointly supported organization and meetings of the FLASH group. Final analysis and publication of results were under the authority of the FLASH group.

Prior Presentation

Presented at the Annual Meeting of the American Society of Clinical Oncology, Chicago, IL, May 29-June 2, 2015; the 20th Congress of the European Hematology Association, Vienna, Austria, June 11-14, 2015; and the 13th International Conference on Malignant Lymphoma, Lugano, Switzerland, June 17-20, 2015.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Thirty-Month Complete Response as a Surrogate End Point in First-line Follicular Lymphoma Therapy: An Individual Patient-level Analysis of Multiple Randomized Trials

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/jco/site/ifc.

Qian Shi
No relationship to disclose

Christopher R. Flowers
Consulting or Advisory Role: OptumRx, Seattle Genetics, Gilead Sciences, AbbVie
Research Funding: Acerta Pharma (Inst), Infinity Pharmaceuticals (Inst), Onyx (Inst), Janssen Pharmaceuticals (Inst), Gilead Sciences (Inst), Celgene (Inst), TG Therapeutics (Inst), Genentech (Inst), Roche (Inst), Pharmacyclics (Inst), AbbVie (Inst)
Travel, Accommodations, Expenses: Celgene, Genentech, Roche

Wolfgang Hiddemann
Consulting or Advisory Role: Roche, Genentech, Celgene, Gilead Sciences, Janssen Pharmaceuticals
Research Funding: Genentech (Inst), Roche (Inst), Janssen Pharmaceuticals (Inst), Gilead Sciences (Inst), Celgene (Inst)

Robert Marcus
Honoraria: Roche
Consulting or Advisory Role: Roche
Speakers' Bureau: Roche
Travel, Accommodations, Expenses: Takeda Pharmaceuticals

Michael Herold
Honoraria: Celgene, Roche
Consulting or Advisory Role: Gilead Sciences

Anton Hagenbeek
No relationship to disclose

Eva Kimby
Honoraria: Roche, Janssen-Cilag, Gilead Sciences
Consulting or Advisory Role: Celgene, Gilead Sciences, Janssen-Cilag, AbbVie, Pfizer
Research Funding: Pfizer (Inst)

Howard Hochster
Consulting or Advisory Role: Bayer AG, Boehringer Ingelheim, Genetech, Amgen, Sirtex Medical, Bristol-Myers Squibb

Umberto Vitolo
Travel, Accommodations, Expenses: Roche, Celgene, Janssen Pharmaceuticals

Bruce A. Peterson
No relationship to disclose

Emmanuel Gyan
Research Funding: Fresenius Kabi (Inst), Mundipharma (Inst), Gilead Sciences (Inst)
Travel, Accommodations, Expenses: Fresenius Kabi, Janssen Pharmaceuticals, Celgene, Gilead Sciences

Michele Ghielmini
Honoraria: Roche, Celgene, Mundipharma, Janssen Pharmaceuticals, Gilead Sciences
Research Funding: Roche Pharma (Schweiz) AG
Travel, Accommodations, Expenses: Celgene, Roche, Genentech

Tina Nielsen
Employment: Roche
Stock or Other Ownership: Genmab, Roche

Sabine De Bedout
Employment: Celgene
Stock or Other Ownership: Celgene

Tommy Fu
Employment: Celgene
Stock or Other Ownership: Celgene

Nancy Valente
Employment: Genentech
Leadership: Genentech
Stock or Other Ownership: Genentech

Nathan H. Fowler
Consulting or Advisory Role: Celgene, Roche, Pharmacyclics, Infinity Pharmaceuticals

Eva Hoster
Honoraria: StiL Forschungs
Research Funding: Roche Pharma (Schweiz) AG, Celgene
Travel, Accommodations, Expenses: Roche Pharma (Schweiz) AG, Celgene, F. Hoffman La-Roche, StiL Forschungs

Marco Ladetto
Research Funding: Amgen (Inst), Celgene (Inst), Gilead Sciences (Inst), Mundipharma (Inst), Novartis (Inst), Johnson & Johnson (Inst), Roche (Inst), Roche Molecular Systems (Inst), Spectrum Pharmaceuticals (Inst), Takeda Pharmaceuticals (Inst)
Travel, Accommodations, Expenses: Amgen, Celgene, Gilead Sciences, Mundipharma, Novartis, Johnson & Johnson, Roche, Roche Molecular Systems, Spectrum Pharmaceuticals, Takeda Pharmaceuticals

Franck Morschhauser
Honoraria: Genentech, Roche, Celgene
Consulting or Advisory Role: Gilead Sciences
Speakers' Bureau: Celgene, Janssen Pharmaceuticals
Expert Testimony: Celgene
Travel, Accommodations, Expenses: Celgene

Emanuele Zucca
Honoraria: Roche, Celgene, Janssen Pharmaceuticals
Consulting or Advisory Role: Roche, Celgene, Janssen Pharmaceuticals, Gilead Science, Takeda Pharmaceuticals, AbbVie
Research Funding: Roche (Inst), Mundipharma (Inst), Celgene (Inst)
Travel, Accommodations, Expenses: Roche, Janssen Pharmaceuticals, Gilead Sciences

Gilles Salles
Honoraria: Roche, Genentech, Amgen, Mundipharma, Janssen Pharmaceuticals, Bristol-Myers Squibb, Celgene, Servier, Gilead Sciences
Consulting or Advisory Role: Roche, Genentech, Gilead Sciences, Janssen Pharmaceuticals, Celgene, Novartis, Novimmune
Research Funding: Roche (Inst), Genentech (Inst)
Travel, Accommodations, Expenses: Roche, Genentech

Daniel J. Sargent
Consulting or Advisory Role: AbbVie, Acerta Pharma, ARIAD Pharmaceuticals, Astellas Pharma, AstraZeneca, Medimmune, Biothera, Celldex Therapeutics, Exelixis, Genentech, Incyte, Kyowa Hakko Kirin, Medivation, Merck, Merrimack Pharmaceuticals, Nektar, Novartis, Pharmacyclics, Pique Therapeutics, Spiration, XBiotech
Research Funding: Celgene (Inst), Roche (Inst), Genentech (Inst)
Travel, Accommodations, Expenses: Celgene

Acknowledgment

This work is dedicated to the memory of Daniel J. Sargent after his untimely passing. Dan was one of the world's foremost experts in biostatistics and oncology who brought together disparate investigators and established data sharing across academia and industry internationally. His groundbreaking initiatives of integrating large collections of databases enabled research to answer questions otherwise beyond statistical possibility, to design important new clinical studies, to make regulatory observations, and to set new standards. He pushed these innovations farther to prospectively plan internationally combined analyses that answered questions previously believed to be impossible. The world of oncology statistics and analysis will not be the same without him, but his legacy continues.

We thank all the investigators of the studies included in the analysis and Charles Foussard (Service d'Hématologie Clinique du Centre Hospitalier Universitaire d'Angers, Angers, France), Oliver W. Press (Fred Hutchinson Cancer Research Center, Seattle, WA), Mathias J. Rummel (Justus-Liebig Universität Giessen, Giessen, Germany), Catherine Sebban (Oncologie/Hématologie du Centre Hospitalier Léon Bérard, Lyon, France), Philippe Solal-Céligny (Institut de Cancérologie de l'Ouest, Angers, France), MMRGlobal (Los Angeles, CA), and Astellas Deutschland (Munich, Germany) for providing additional study data. We also thank group statisticians Sami Boussetta (Lymphoma Academic Research Organisation, Centre Hospitalier Lyon-Sud, Pierre-Bénite, France) and Catherine Fortpied (European Organisation for Research and Treatment of Cancer, Brussels, Belgium) as well as Marie-Laure Bravo (Celgene, Boudry, Switzerland), Kenichi Takeshita (Celgene, Summit, NJ), and Jane Huang and Fan Zhang (Genentech, South San Francisco, CA) for helpful discussions; Robert Rydzewski, (Bio Connections, Chicago, IL) for editorial support; and Jean-Pierre Bizzari (Celgene, Summit, NJ) for first proposing the idea that a surrogate end point be developed for follicular lymphoma. Editorial support was provided by Bio Connections, which was funded by Celgene.

Members of the FLASH group are Daniel J. Sargent, chair (Mayo Clinic, Rochester, MN), Christopher R. Flowers (Emory University, Atlanta, GA), Nathan H. Fowler (University of Texas MD Anderson Cancer Center, Houston, TX), Anton Hagenbeek (Academic Medical Center, Amsterdam, the Netherlands), Michael Herold (HELIOS Kliniken, Erfurt, Germany), Wolfgang Hiddemann (Ludwig-Maximilians University Hospital, Munich, Germany), Eva Kimby (Karolinska Institute, Stockholm, Sweden), Marco Ladetto (Azienda Ospedaliera SS Antonio e Biagio e Cesare Arrigo, Alessandria, Italy), Robert Marcus (Addenbrooke's Hospital, Cambridge, United Kingdom), Franck Morschhauser (Service Université de Lille 2, Lille, France), Gilles Salles (Université Claude Bernard, Pierre Bénite, France), Umberto Vitolo (Azienda Ospedaliero-Universitaria Città della Salute e della Scienza di Torino, Turin, Italy), and Emanuele Zucca (Swiss Group for Clinical Cancer Research, Bern, Switzerland).

Appendix

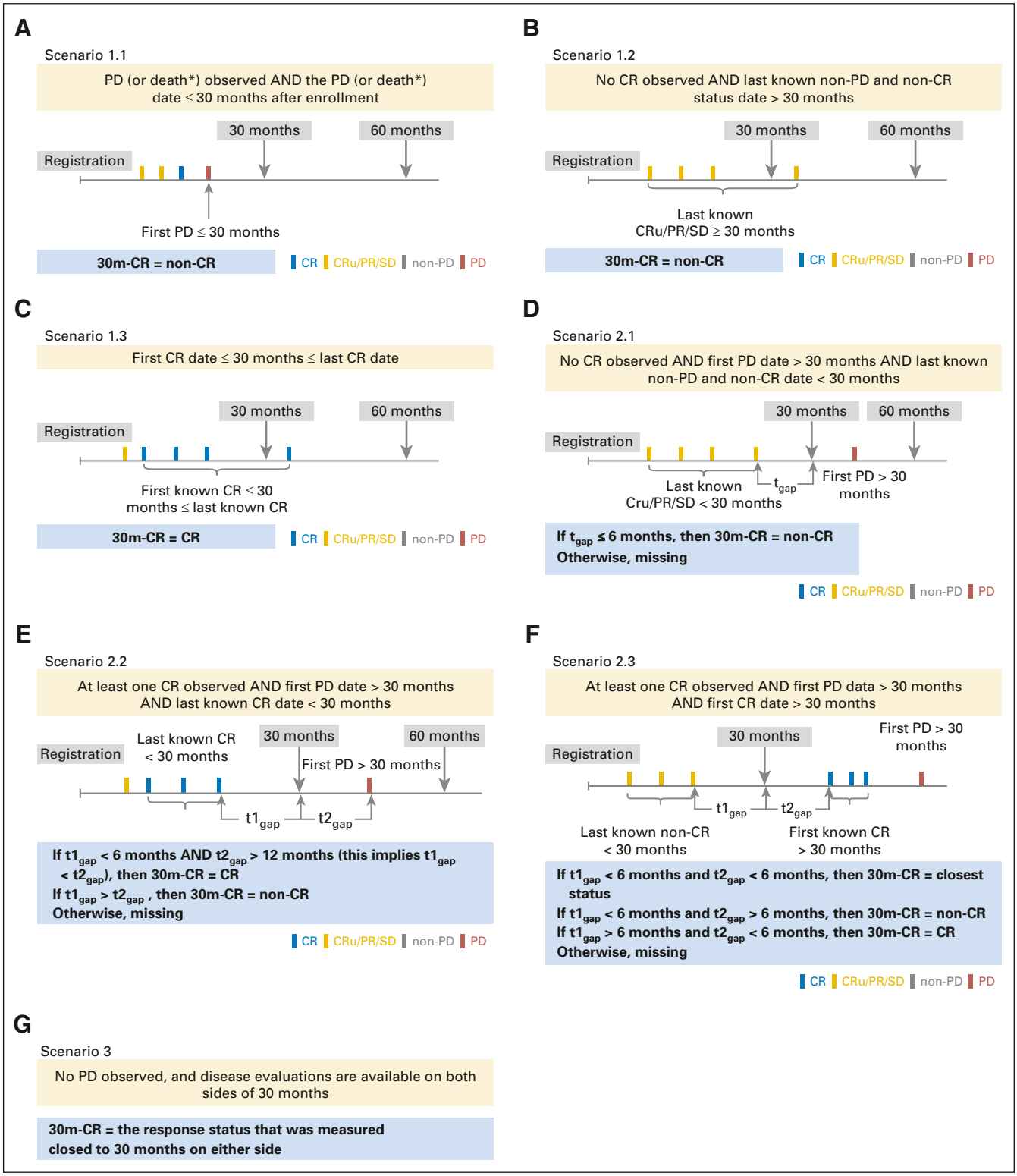

Fig A1. Secondary calculation rules for complete response (CR) rate at 30 months (CR30) after initiation of induction treatment. (A to G) The determination algorithm (stated in the blue boxes) of 30-month CR status (30m-CR) is illustrated according to various scenarios. CR, complete response; CRu, unconfirmed complete response; PD, progressive disease; PR, partial response; SD, stable disease; t_{gap}, gap between two time points.

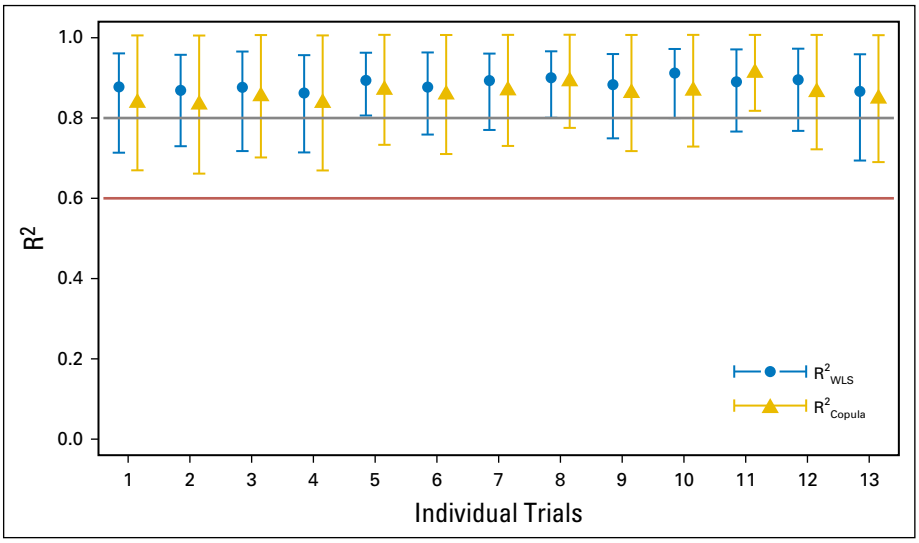

Fig A2. Re-estimating R^2_{WLS} and R^2_{Copula} by leaving one trial out at a time (complete response rate at 30 months [CR30] based on primary imputation rules). For each labeled trial, R^2_{WLS} and R^2_{Copula} were estimated by excluding the labeled trial. The circles and triangles indicate the point estimates for R^2_{WLS} and R^2_{Copula}, respectively. The solid bars are the 95% CIs estimated on the basis of the quantiles of bootstrap samples for R^2_{WLS} and R^2_{Copula}, respectively. Gray and red lines indicate the cutoff points stated in the surrogacy evaluation criteria for point estimate (≥ 0.8) and lower bound of 95% CI (> 0.6), respectively. The numbers sequentially correspond to the following trials: Cancer and Leukemia Group B (CALGB) 7951, Eastern Cooperative Oncology Group (ECOG) E1496, European Organization for Research and Treatment of Cancer (EORTC) 20921, Phase III Trial of Favid and Gmcsf [granulocyte-macrophage colony-stimulating factor] vs Placebo and Gmcsf Following Rituximab (FavId06), French Acute Leukemia and Blood Diseases West-East Group (GOELAMS) 064, M39021, M39023, ML16865, ML17638, Advanced Ovarian Cancer Following Response on Front-Line Platinum-Based Chemotherapy (PRIMA), Swiss Group for Clinical Cancer Research (SAKK) 35/98, German Low-Grade Lymphoma Study Group (GLSG) 2000, and GLSG1996. Copula, bivariate Plackett copula model; WLS, weighted least squares.