Progression-Free Survival as a Surrogate End Point for Overall Survival in First-Line Diffuse Large B-Cell Lymphoma: An Individual Patient–Level Analysis of Multiple Randomized Trials (SEAL)

Qian Shi, Norbert Schmitz, Fang-Shu Ou, Jesse G. Dixon, David Cunningham, Michael Pfundtshuh, John F. Seymour, Ulrich Jaeger, Thomas M. Habermann, Corinne Haionn, Hervé Tilly, Hervé Ghesquieres, Francesco Merli, Marita Ziepert, Raoul Herbrecht, Jocelyne Flament, Tommy Fu, Bertrand Coiffier, and Christopher R. Flowers

Purpose
Overall survival (OS) is the definitive and best-established primary efficacy end point to evaluate diffuse large B-cell lymphoma (DLBCL) therapies, but it requires prolonged follow-up. An earlier end point assessed post-treatment would expedite clinical trial conduct and accelerate patient access to effective new therapies. Our objective was to formally evaluate progression-free survival (PFS) and PFS at 24 months (PFS24) as surrogate end points for OS in first-line DLBCL.

Patients and Methods
Individual patient data were analyzed from 7,507 patients from 13 multicenter randomized controlled trials of active treatment in previously untreated DLBCL, published after 2002, with sufficient PFS data to predict treatment effects on OS. Trial-level surrogacy examining the correlation of treatment effect estimates of PFS/PFS24 and OS was evaluated using both linear regression (RWLS) and Copula bivari (Rcopula) models. Prespecified criteria for surrogacy required either RWLS or Rcopula ≥ 0.80 and neither < 0.7, with lower-bound 95% CI > 0.60.

Results
Trial-level surrogacy for PFS was strong (RWLS = 0.83; Rcopula = 0.85) and met the predefined criteria for surrogacy. At the patient level, PFS strongly correlated with OS. The surrogate threshold effect had a hazard ratio of 0.89. Surrogacy was consistent across comparisons with or without rituximab and with rituximab maintenance trials. Trial-level surrogacy for PFS24 was relatively strong (RWLS = 0.77; Rcopula = 0.78) but did not meet prespecified criteria. At the patient level, PFS24 significantly correlated with OS. The surrogate threshold effect had an odds ratio of 1.51.

Conclusion
This large pooled analysis of individual patient data supports PFS as a surrogate end point for OS in future randomized controlled trials evaluating chemoimmunotherapy in DLBCL. Use of this end point may expedite therapeutic development with the intent of bringing novel therapies to this patient population years before OS results are mature.

J Clin Oncol 36:2593-2602. © 2018 by American Society of Clinical Oncology

INTRODUCTION
Diffuse large B-cell lymphoma (DLBCL) accounts for approximately 32% of non-Hodgkin lymphoma (NHL) diagnoses.1 The addition of rituximab to cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP) as initial therapy for DLBCL was one of the key contributors to improved 5-year standardized cancer-specific survival from 52% in 1999 to 66% in 20052 and led to R-CHOP becoming the most common first-line treatment. Relative 10-year survival for patients with DLBCL is approximately 50%,3 illustrating the persisting need for improved therapies. However, no regimen has demonstrated a survival benefit compared with R-CHOP in a randomized trial.3-5 Approximately 10% to 15% of patients with DLBCL have disease that is refractory to initial therapy, and another 20% to 25% who experience relapse after an initial
response to treatment will experience worse outcomes.6,7 Ten-year DLBCL-specific survival is considerably dependent on the extent of disease and prognostic score and has been shown to be 80% for patients with low-risk, 60% for those with intermediate-risk, and 36% for those with high-risk disease.8 Genomic tools for risk stratification and molecular surveillance of DLBCL have been presented,9,10 but neither these tools nor novel agents have improved patient outcomes.

Overall survival (OS) is the gold-standard primary efficacy end point for clinical trials of antilymphoma treatment. On the basis of studies evaluated here, it may take >8 years to achieve median OS for first-line DLBCL and, depending on accrual time, approximately 10 years to complete phase III trials using OS as an end point. For example, a randomized controlled trial (RCT) comparing R-CHOP with an alternative dose-intensive regimen (CALGB [Cancer and Leukemia Group B]/Alliance 50303) opened in 2005 and reported initial findings in December 2016.5 Results from previous studies in DLBCL and other aggressive types of NHL suggest that progression-free survival (PFS) and short-term outcomes, including event-free survival and PFS at 0.5, 2, or 3 years, should be explored as candidates for an earlier surrogate end point for OS.11-14 However, these studies were based on suboptimal data and/or validation methods.

We established the Surrogate Endpoints for Aggressive Lymphoma (SEAL) international collaboration to construct a large database integrating individual patient data (IPD) from completed, published, multicenter RCTs in DLBCL to evaluate potential surrogate end points for OS in DLBCL trials and support continuous translational research (eg, prognostic analyses, risk classifications, subgroup analyses).15 Several studies have been successful in using metadatabases to provide a strong foundation and methodology for establishing statistical criteria that assess clinically meaningful relationships between end points and surrogates. These studies combined IPD from multiple studies to identify earlier surrogate end points for adjuvant treatment in colorectal cancer (ACCENT [Adjuvant Colon Cancer End Points]);16 metastatic colorectal cancer (ARCAD [Analysis and Research in Cancers of the Digestive System]);17 and follicular lymphoma (FLASH [Follicular Lymphoma Analysis of Surrogate Hypothesis]).18 The specific objectives of the SEAL study are to determine trial-level (primary) and patient-level (secondary) correlations between two surrogate end point candidates, PFS and PFS at 24 months (PFS24), with OS after first-line treatment of DLBCL using well-established statistical methods on IPD. Here we report results of an analysis of data from 13 studies in a total of 7,507 patients.

PATIENTS AND METHODS

**Trial Selection and Comparison Definition**

On June 30, 2015, we used the search terms DLBCL, NHL, and aggressive or advanced or diffuse histiocytic in the title and keywords up-front or first-line or untreated or newly diagnosed or initial/primary/treatment/therapy or no prior therapy or naive to conduct a comprehensive search of published DLBCL clinical studies in the Medline database maintained by the US National Library of Medicine. To be included, studies were required to be multicenter RCTs published in English from 2002 to October 13, 2015, and designed to evaluate active treatment in ≥100 adult (no pediatric) patients with previously untreated DLBCL or aggressive NHL defined by WHO/Revised European American Lymphoma classification. Studies were excluded if they were reviews; if they enrolled patients with early-stage (I or II) disease only, low-grade or indolent or HIV-related lymphoma, or relapsed or refractory disease; or if the intervention focused on salvage treatment or supportive care, such as growth factors, palliative care, quality of life, or health economics.

The sponsors of all identified studies meeting the prespecified inclusion criteria were contacted to determine their interest in confidentially sharing IPD. Studies unable to transfer IPD before December 2016 were excluded for this analysis. All exclusions were based only on data quality and availability and were determined before any statistical analysis of the end points.

As of August 2016, 13 of 14 studies collected in the SEAL database met the inclusion criteria. Among those studies, there were 10 studies comparing induction treatments only, two studies comparing both induction and maintenance treatments, and one study comparing maintenance treatments only. One study (RICOVER60 [Rituximab With CHOP Over Age 60 Years]) included three experimental arms compared with one control arm. The meta-analytic unit for surrogacy estimation was predefined as the comparison between two arms (experimental v control) nested within trials. The induction and maintenance comparisons within the same study were treated as two comparison units. As such, a total of 17 comparison units were predefined (Table 1).

**Surrogate End Point Candidates**

Potential surrogate candidates for OS included PFS and PFS24. PFS was a time-to-event end point defined as the time from initiation of induction treatment to the earliest occurrence of progressive disease, relapse, or death resulting from any cause. Living patients without documented disease progression were censored on the date of their last disease evaluation. PFS24 was a binary end point where patients who were alive and in disease-free status up to 24 months after initiation of induction treatment were considered a success. The primary end point and surrogate candidates were derived according to consistent calculation rules across studies.

**Statistical Methods**

**True end point.** The primary clinical end point was OS, defined as time from initiation of induction treatment to death resulting from any cause. Living patients were censored on the date when they were last documented as alive.

**General statistical methods.** Within-trial treatment effects for time-to-event end points OS and PFS were quantified using hazard ratios (HRs). Within-trial treatment effects for the binary end point PFS24 were quantified using odds ratios (ORs), and 95% CIs were calculated for each. Analyses were performed by using SAS software (version 9.4; SAS Institute, Cary, NC) and R software (version 2.14; R Foundation for Statistical Computing, Vienna, Austria).

**Surrogacy evaluation.** The primary surrogacy evaluation method was trial-level surrogacy, which measured how precisely treatment effect on the true end point may be predicted based on observed treatment effects on the surrogate end point. At the trial level, two commonly used trial-level surrogacy measures were considered: Copula bivariate (\(R_{\text{copula}}^2\))15,36 and linear regression based on weighted least squares regression method (\(R_{\text{OLS}}^2\)),19 where \(R_{\text{copula}}^2\) takes into account patient-level correlation between the two end points and \(R_{\text{OLS}}^2\) does not. After publication of surrogacy qualification criteria established in the follicular lymphoma setting,18 the predefined rule for declaring trial-level surrogacy required either \(R_{\text{OLS}}^2 \geq 0.80\), with the lower boundary of a 95% CI of >0.6, and neither estimate <0.7.

Supplemental trial-level surrogacy measures included the surrogacy threshold effect,20 the minimum treatment effect on the surrogate required to confidently predict a significant treatment effect on OS in a future trial. For patient-level surrogacy, the correlations between the candidate surrogate end point and OS were quantified through rank correlation coefficient (\(p\) for PFS via the biviable Clayton Copula model19; for PFS24, the global OR for comparing the odds of remaining alive beyond a particular time point between
patients with different PFS24 status was estimated through the bivariable Plackett Copula model over the entire duration of OS. Patient-level correlation was considered a supportive but not sufficient condition for surrogacy validation. A closer to 1.0 indicated a stronger correlation. For the global OR, a 95% CI excluding 1 indicated significant individual-level correlation.

**Sensitivity and Subpopulation Analysis**

Leave-one-out cross-validation, which compares the predicted log (HR) values with the observed values on OS based on the estimated trial-level model, leaving one trial out at a time, assessed the predictive performance of the regression model. Leave-one-out estimation, which re-estimated $R^2$ when one trial was excluded at a time, was performed to identify potentially influential trials. Surrogacy was further examined within subpopulations defined by treatment type (ie, comparisons with rituximab, induction comparisons only).

**RESULTS**

**Patient Characteristics**

Of a total of 267 references individually examined, 14 studies were included in the SEAL database and 13 studies met all inclusion criteria (Fig 1). Table 1 summarizes trial-level characteristics of the 13 studies included, involving IPD from 7,507 patients and 17 two-arm comparisons, of which 14 compared...
induction regimens and three compared maintenance regimens. Fifteen comparisons involved rituximab.

Baseline characteristics were well balanced between arms (Table 2), with a median age of 62.8 years; 42% of patients were age < 60 years, 32% were age 60 to 69 years, and 26% were ≥ 70 years; 54% were men, and 14% had Eastern Cooperative Oncology Group performance status ≥ 2. One third of patients (33%) had International Prognostic Index (IPI) scores of 0 to 1; 25%, IPI 2; 24%, IPI 3; and 17%, IPI 4 to 5. Ann Arbor stage was I/II in 36% of patients, III in 24%, and IV in 40%. Overall, the median follow-up time for OS was 52 months. The distributions of PFS and OS in overall population pooling all patients across studies are shown in Appendix Figures A1A and A1B (online only).

Surrogacy was consistent across comparisons with and without rituximab induction and with rituximab maintenance.

Leave-one-out cross-validation demonstrated consistency between observed and predicted OS treatment effects for each comparison unit on the basis of PFS (Fig 2B), except when leaving one comparison (PIX203) out. This comparison was also shown to be a high influence outlier when evaluating the re-estimated $R^2$ when one comparison at a time was excluded (Fig 2C). Exclusion of this comparison yielded an $R_{WLS}$ of 0.87 (95% CI, 0.70 to 0.95) and $R_{Copula}^2$ of 0.88 (95% CI, 0.76 to 0.99).

**PFS24: Trial- and Patient-Level Surrogacy Measures**

As summarized in Table 3 and Figure 3A, trial-level surrogacy for PFS24 ($R_{WLS} = 0.77$; 95% CI, 0.51 to 0.92; $R_{Copula}^2 = 0.78$; 95% CI, 0.59 to 0.96) was slightly less robust than PFS and did not meet the prespecified surrogacy qualification criteria. At the patient level, the global OR was 61.1 (95% CI, 52.6 to 69.6), which indicates substantially higher odds of remaining alive beyond a particular time point for patients who were alive and disease free at 24 months after initiation of induction treatment (Table 3). The surrogate threshold effect had an OR of 1.51, which indicated that an observed OR ≥ 1.51 for PFS24 would predict a significant treatment effect on OS in a future trial. Surrogacy performance improved when restricted to

---

**References identified (N = 267)**

- Excluded/did not meet ≥ one inclusion criterion (n = 251)
- Considered eligible; owners contacted (n = 16)
- Excluded by steering committee: end points difficult to interpret (n = 1)
- Did not provide IPD (n = 1)
- Provided IPD; included in SEAL database (n = 14)
- Excluded from surrogacy analysis; (n = 1) enrolled early-stage (I or II) disease only

**Table 2. Patient Characteristics Based on Trials Included in the Primary Analysis (n = 13)**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Control (n = 3,450)</th>
<th>Experimental (n = 4,057)</th>
<th>Total (n = 7,507)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (continuous), years</td>
<td>Mean 58.6</td>
<td>60.0</td>
<td>59.3</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>SD</td>
<td>14.7</td>
<td>14.2</td>
<td>14.4</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>62.0</td>
<td>630</td>
<td>62.8</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>18.0-89.6</td>
<td>18.0-92.2</td>
<td>18.0-92.2</td>
<td></td>
</tr>
<tr>
<td>Age (categorical), years</td>
<td>&lt; 60</td>
<td>1,566 (45)</td>
<td>1,562 (39)</td>
<td>3,128 (42)</td>
</tr>
<tr>
<td>60-69</td>
<td>1,034 (30)</td>
<td>1,386 (34)</td>
<td>2,420 (32)</td>
<td></td>
</tr>
<tr>
<td>≥ 70</td>
<td>850 (25)</td>
<td>1,109 (27)</td>
<td>1,959 (26)</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>Female</td>
<td>1,580 (46)</td>
<td>1,896 (47)</td>
<td>3,476 (46)</td>
</tr>
<tr>
<td>Male</td>
<td>1,870 (54)</td>
<td>2,161 (53)</td>
<td>4,031 (54)</td>
<td></td>
</tr>
<tr>
<td>ECOG PS</td>
<td></td>
<td></td>
<td></td>
<td>.1988</td>
</tr>
<tr>
<td>Missing</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1,627 (47)</td>
<td>1,837 (45)</td>
<td>3,464 (46)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1,328 (38)</td>
<td>1,641 (40)</td>
<td>2,969 (40)</td>
<td></td>
</tr>
<tr>
<td>≥ 2</td>
<td>492 (14)</td>
<td>578 (14)</td>
<td>1,070 (14)</td>
<td></td>
</tr>
<tr>
<td>IPI score*</td>
<td>Missing</td>
<td>393</td>
<td>394</td>
<td>777</td>
</tr>
<tr>
<td>0-1</td>
<td>1,022 (33)</td>
<td>1,217 (33)</td>
<td>2,239 (33)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>734 (24)</td>
<td>968 (26)</td>
<td>1,702 (25)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>768 (25)</td>
<td>878 (24)</td>
<td>1,646 (24)</td>
<td></td>
</tr>
<tr>
<td>4-5</td>
<td>533 (17)</td>
<td>610 (17)</td>
<td>1,143 (17)</td>
<td></td>
</tr>
<tr>
<td>Ann Arbor stage</td>
<td>Missing</td>
<td>14</td>
<td>9</td>
<td>23</td>
</tr>
<tr>
<td>I/II</td>
<td>1,223 (35)</td>
<td>1,492 (37)</td>
<td>2,715 (36)</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>1,767 (53)</td>
<td>1,022 (26)</td>
<td>1,809 (24)</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>1,426 (41)</td>
<td>1,534 (38)</td>
<td>2,960 (40)</td>
<td></td>
</tr>
</tbody>
</table>

*Percentages were calculated within each group after subtraction of missing IPI score data.

**Abbreviations:** ECOG PS, Eastern Cooperative Oncology Group performance status; IPI, International Prognostic Index; SD, standard deviation.

---

Fig 1. Flowchart of study selection. Abbreviation: IPD, individual patient data.
induction comparisons. Sensitivity analyses showed similar results (Figs 3B and 3C).

This pooled analysis compiled IPD from 13 DLBCL RCTs published from 2002 to 2015. The acquisition and integration of IPD from these trials were conducted by an independent data center at Mayo Clinic according to predefined end point derivations and statistical plans. The SEAL group, which conducted these analyses, included involvement from international lymphoma experts and biostatisticians who had previously developed these statistical approaches in consultation with regulatory agencies from the United States and Europe. To our knowledge, this is the first analysis based on integrated IPD from RCTs involving patients with DLBCL. Unlike literature-based meta-analyses, use of IPD ensured the consistent determination of end points and the consistent interpretation of within-trial treatment effects across all studies. Our analysis demonstrates that treatment effect on PFS is a strong predictor of treatment effect on OS. These results were consistent across different surrogate estimation methods and sensitivity analyses. The strong association between PFS and OS was maintained irrespective of the inclusion of rituximab in the induction and/or maintenance regimen. Together, these results indicate that PFS serves as a strong surrogate for OS in trials evaluating first-line therapy for DLBCL.

DLBCL is the most prevalent aggressive form of NHL and a heterogeneous disease with subtypes distinguishable by molecular, immunophenotypic, morphologic, and clinical characteristics. In the prerituximab era, the chemotherapy combination CHOP was established as standard first-line treatment of DLBCL based on results from a randomized study comparing CHOP with three other more aggressive combination chemotherapy regimens. Integration of the monoclonal anti-CD20 antibody rituximab into chemotherapy regimens was one of the most important advances in the treatment of DLBCL based on data first published in 2002 and confirmed in subsequent trials. Despite promising data from early-phase clinical trials, since 2002, no regimen has demonstrated significant clinical benefit over R-CHOP in a randomized controlled clinical trial. Surrogate end points such as PFS are needed to reduce the evaluation time for new agents and regimens so that effective first-line regimens and strategies can be delivered in clinical practice sooner, and ineffective regimens can be abandoned without prolonged evaluation.

Current approaches for improving outcomes in DLBCL focus on addressing therapeutic targets present in nearly all patients with DLBCL, such as CD19 and CD79b, and characterizing the genetic mutations and pathways involved in DLBCL along with their functional roles to define risk stratification models and identify new therapeutic targets. Clinically, patients with DLBCL who achieve event-free survival at 24 months after initiation of therapy experience survival similar to that of matched controls, whereas those who do not have poor expected survival. Targeting poor-risk patients identified by clinical or genetic factors has become a priority for defining patient groups, where improving upon outcomes with R-CHOP is needed and feasible in a suitable timeframe for drug development. Recognition and continued investigation into the complexity and heterogeneity of DLBCL led to classifications based on immunohistochemistry and gene expression profiling that distinguish two major subtypes of DLBCL: germinal center B cell–like DLBCL and activated B cell–like DLBCL, with the latter type being associated with worse prognosis in some but not all studies. Fluorescence in situ hybridization and immunohistochemistry also differentiate patients with DLBCL who are positive for MYC and BCL2 and/or BCL6; these patients experience worse outcomes when treated with R-CHOP.

Precision medicine approaches that perform DLBCL subtyping in all patients and then add a novel agent to R-CHOP therapy for a poor-risk subgroup have been projected to provide benefits that outweigh their costs, but at present, such strategies have failed to demonstrate clinical benefit. Three RCTs demonstrated that substituting bortezomib for vincristine or adding bortezomib to R-CHOP did not significantly improve outcomes. A randomized trial substituting obinutuzumab for rituximab demonstrated that this regimen did not significantly improve investigator-assessed PFS compared with R-CHOP. Likewise, altering the administration of R-CHOP to every 14 days instead of every 21 days increased toxicity without providing benefits. Failure of these trials has been attributed to selection bias, leading to prospectively enrolled patients with DLBCL who had more favorable survival than was expected with R-CHOP based on historical controls, and limited additive benefits of these approaches. These trials also reinforce the use of PFS in the clinical research community as an acceptable surrogate for OS. Novel agents and approaches are emerging from early phase I/II clinical trials that may improve survival.

### Table 3. Trial- and Patient-Level Surrogacy Measures

<table>
<thead>
<tr>
<th>Outcome</th>
<th>No. of Trials (No. of patients)</th>
<th>Trial-Level Surrogacy</th>
<th>Patient-Level Surrogacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>PFS (time-to-event end point)</td>
<td></td>
<td>[R^2_{WLS} \ (95% CI)]</td>
<td>[R^2_{adj} \ (95% CI)]</td>
</tr>
<tr>
<td>Overall</td>
<td>17 (7,507)</td>
<td>0.83 (0.57 to 0.94)</td>
<td>0.85 (0.73 to 0.98)</td>
</tr>
<tr>
<td>Comparisons with rituximab</td>
<td>15 (7,001)</td>
<td>0.81 (0.55 to 0.93)</td>
<td>0.86 (0.72 to 0.99)</td>
</tr>
<tr>
<td>Induction comparisons only</td>
<td>14 (6,826)</td>
<td>0.80 (0.53 to 0.94)</td>
<td>0.82 (0.66 to 0.99)</td>
</tr>
<tr>
<td>PFS24 (binary end point)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>17 (6,882)</td>
<td>0.77 (0.51 to 0.92)</td>
<td>0.78 (0.59 to 0.96)</td>
</tr>
<tr>
<td>Comparisons with rituximab</td>
<td>15 (6,180)</td>
<td>0.75 (0.45 to 0.91)</td>
<td>0.78 (0.59 to 0.98)</td>
</tr>
<tr>
<td>Induction comparisons only</td>
<td>14 (6,047)</td>
<td>0.80 (0.53 to 0.94)</td>
<td>0.82 (0.66 to 0.99)</td>
</tr>
</tbody>
</table>

Abbreviations: OR, odds ratio; PFS, progression-free survival; PFS24, PFS at 24 hours; WLS, weighted least squares. *\[\rho\] for PFS and global OR for PFS24. \[\rho\] is the rank correlation coefficient quantifying the individual-level correlation between PFS and OS; \[\rho\] closer to 1.0 indicated a stronger correlation. Global OR quantifies the individual-level correlation between PFS24 and OS. A 95% CI excluding 1 indicated significant individual-level correlation.
Fig 2. (A) Trial-level treatment effect correlation between progression-free survival (PFS) and overall survival (OS). Circle/triangle size is proportional to sample size. Solid line indicates fitted weighted least squares (WLS) regression line; gray dashed lines indicate 95% prediction limits. (B) PFS surrogacy sensitivity analysis: leave-one-out cross-validation. For each comparison, the open circle/triangle is the predicted log (hazard ratio [HR]) on OS based on the estimated WLS regression line at trial level, removing the comparison listed; the horizontal bars indicate 95% prediction interval. (C) PFS surrogacy sensitivity analysis: leave-one-out re-estimation. For each labeled comparison, $R^2_{WLS}$ and $R^2_{Copula}$ were estimated by excluding the labeled comparison; vertical bars indicate 95% CI. ANZINTER, Intergruppo Italiano Linfomi; MegaCHOEP, cyclophosphamide, doxorubicin, vincristine, and prednisone plus etoposide; MInT, MabThera International Trial Group; NHL, non-Hodgkin lymphoma; PIX, pixantrone; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone.
Fig 3. (A) Trial-level treatment effect correlation between progression-free survival at 24 months (PFS24) and overall survival (OS). Circle/triangle size is proportional to sample size. Solid line indicates fitted weighted least squares (WLS) regression line; gray dashed lines indicate 95% prediction limits. (B) PFS24 surrogacy sensitivity analysis: leave-one-out cross-validation. For each comparison, the open circle/triangle is the predicted log (hazard ratio [HR]) on OS based on the estimated WLS regression line at trial level, removing the comparison listed; horizontal bars indicate 95% prediction interval. (C) PFS24 surrogacy sensitivity analysis: leave-one-out re-estimation. For each labeled comparison, $R^2_{\text{WLS}}$ and $R^2_{\text{Copula}}$ were estimated by excluding the labeled comparison; vertical bars indicate 95% CI. ANZINTER, Intergruppo Italiano Linfomi; MegaCHOEP, cyclophosphamide, doxorubicin, vincristine, and prednisone plus etoposide; MInT, MabThera International Trial Group; NHL, non-Hodgkin lymphoma; PIX, pixantrone; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone.
for all patients with DLBCL, with particular attention to poor-risk groups that may benefit more from a selected therapy. Randomized studies are ongoing to demonstrate the potential benefit of these approaches over R-CHOP. These trials and future studies should examine design strategies to reduce selection bias and include surrogate end points (such as PFS) to improve the likelihood of success, along with reducing the time for evaluating the clinical benefit of interventions.

Future research examining PFS24 and new candidates for surrogacy should be carried out as the SEAL collaboration continues to actively pursue additional trials and data. For emerging novel agents substituted in or combined with R-CHOP, PFS can be an appropriate surrogate end point when the expected mechanism for improving OS is long-term disease control. For agents expected to improve OS in settings where DLBCL progression is expected to occur on treatment surrogacy, re-evaluation may be required. A majority of clinical trials included in our analysis focused on the induction setting. As a result, a subgroup analysis of PFS surrogacy in the maintenance setting was not possible. As more maintenance studies (eg, REMARC [Study of Lenalidomide Maintenance Versus Placebo in Responding Elderly Patients With DLBCL and Treated With R-CHOP] NCT01122472) become available to SEAL, the surrogacy of PFS and PFS24 could be specifically tested in the maintenance setting. Additional surrogate end point candidates in DLBCL, such as response-based end points and event-free survival, should be evaluated as additional data from current SEAL studies and new studies are added.

In the first-line setting, DLBCL treatment is administered with the intent to achieve long-term disease control and ultimately a cure. Although R-CHOP provides significant clinical benefits to patients, these advantages challenge the conduct of future trials of newer therapeutic options. More efficient approaches are needed to promote the development of novel therapies in a timely manner, provide patients earlier access to more effective therapeutic options, and more rapidly evaluate and discard approaches that do not have a clinically significant impact. Establishing surrogate RCT end points that are measured earlier than current approaches can facilitate future drug development. The results of the SEAL analyses have demonstrated that treatment effects on PFS strongly predicts for treatment effects on OS. Although PFS24 was significantly correlated with longer OS at the patient-level analysis, it did not meet the trial-level prespecified surrogacy qualification criteria. Future analyses with additional trials should re-evaluate the role of PFS24 as a surrogate end point, focusing on the trial-level correlation with OS. In conclusion, this pooled analysis of IPD from randomized clinical trials of first-line treatments for patients with DLBCL demonstrates that the primary surrogate candidate, PFS, met the qualification criteria to be a robust surrogate end point for OS. These results support the use of PFS as an appropriate primary end point in future studies evaluating chemoimmunotherapy in patients with DLBCL.
39. Younes A, Thieblemont C, Morschhauser F, et al: Combination of ibritinib with rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP) for treatment-naive patients with...


Affiliations

Qian Shi, Fang-Shu Ou, Jesse G. Dixon, and Thomas M. Habermann, Mayo Clinic, Rochester, MN; Norbert Schmitz, University of Muenster, Muenster; Michael Pfreundschuh, Universitaet des Saarlandes, Homburg; Marita Ziepert, University of Leipzig, Leipzig, Germany; David Cunningham, The Royal Marsden Hospital, Surrey, United Kingdom; John F. Seymour, Peter MacCallum Cancer Centre and University of Melbourne, Melbourne, Victoria, Australia; Ulrich Jaeger, Medical University of Vienna, Vienna, Austria; Corinne Haouin, Hôpital Henri Mondor, Assistance Publique–Hôpitaux de Paris, Créteil; Hervé Tilly, Institut National de la Santé et de la Recherche Médicale U1245, Université de Rouen, Rouen; Hervé Ghyseliers and Bertrand Coifier, Centre Hospitalier Lyon-Sud, Pierre-Benite; Raoul Herbrecht, Hôpital de Hautepierre, Strasbourg, France; Francesco Merli, Azienda Unità Sanitaria Locale–Istituto di Ricovero e Cura a Carattere Scientifico, Reggio Emilia, Italy; Jocelyne Flament, Celgene, Boudry, Switzerland; Tommy Fu, Celgene, Summit, NJ; and Christopher R. Flowers, Emory University, Atlanta, GA.

Support

Supported by grants from Celgene.

Prior Presentation

Presented in part at the 58th Annual Meeting of the American Society of Hematology, San Diego, CA, December 3-6, 2016.
AUTHORS’ DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Progression-Free Survival as a Surrogate End Point for Overall Survival in First-Line Diffuse Large B-Cell Lymphoma: An Individual Patient–Level Analysis of Multiple Randomized Trials (SEAL)

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

Norbert Schmitz
Stock or Other Ownership: Celgene
Honoraria: Riemser, Takeda Pharmaceuticals, Novartis, Kite Pharma, Janssen Oncology, Hexal
Consulting or Advisory Role: Hexal, Kite Pharma, Riemser
Research Funding: Janssen Oncology
Travel, Accommodations, Expenses: Roche, Riemser, Novartis, Kite Pharma, Janssen Oncology

Fang-Shu Ou
No relationship to disclose

Jesse G. Dixon
No relationship to disclose

David Cunningham
Research Funding: Amgen (Inst), AstraZeneca (Inst), Bayer HealthCare Pharmaceuticals (Inst), Celgene (Inst), Merrimack (Inst), MedImmune (Inst), Merck Serono (Inst), Sanofi (Inst)

Michael Pfreundschuh
Honoraria: Roche, Takeda Pharmaceuticals
Consulting or Advisory Role: Amgen, Boehringer Ingelheim, Celgene, Roche, Spectrum Pharmaceuticals, Novartis, Bayer HealthCare Pharmaceuticals, Bristol-Myers Squibb, Janssen, Sandoz, Mundipharma, MorphoSys, Seattle Genetics, Merck Sharp & Dohme, Kite Pharma
Research Funding: Amgen, Roche, Spectrum Pharmaceuticals
Travel, Accommodations, Expenses: Amgen, Roche, Janssen

John F. Seymour
Honoraria: AbbVie, Celgene, Janssen, Roche
Consulting or Advisory Role: AbbVie, Celgene, Janssen, Roche
Speakers’ Bureau: AbbVie, Roche
Research Funding: AbbVie, Janssen, Roche
Travel, Accommodations, Expenses: Roche, AbbVie, Celgene

Ulrich Jaeger
Honoraria: Amgen, Celgene, AbbVie, Roche, Novartis, Gilead Sciences
Consulting or Advisory Role: Amgen, Celgene, AbbVie, Roche, Novartis, Gilead Sciences
Research Funding: Novartis (Inst), Gilead Sciences (Inst), Roche (Inst),

Thomas M. Habermann
No relationship to disclose

Corinne Haioun
Consulting or Advisory Role: Roche, Celgene, Janssen-Cilag, Gilead Sciences, Amgen, Novartis, Takeda Pharmaceuticals
Travel, Accommodations, Expenses: Roche, Amgen,

Hervé Tilly
Honoraria: Celgene, Roche/Genentech, Servier
Consulting or Advisory Role: Roche, AstraZeneca, Karyopharm Therapeutics, Bristol-Myers Squibb
Research Funding: Celgene (Inst)
Travel, Accommodations, Expenses: Roche

Hervé Ghesquieres
Honoraria: Gilead Sciences, Sanofi
Consulting or Advisory Role: Celgene

Francesco Merli
Honoraria: Roche
Consulting or Advisory Role: Roche, Celgene
Research Funding: Roche
Travel, Accommodations, Expenses: Roche

Marita Ziepert
Research Funding: Asklepios Kliniken Hamburg (Inst)

Raoul Herbrecht
Consulting or Advisory Role: Astellas Pharma, Basilea, Gilead Sciences, Merck Sharp & Dohme, Novartis, Pfizer
Speakers’ Bureau: Pfizer, Gilead Sciences, Basilea
Research Funding: Cell Therapeutics, Pfizer
Travel, Accommodations, Expenses: Basilea, Celgene, Gilead Sciences, Merck Sharp & Dohme, Pfizer

Jocelyne Flament
Employment: Celgene
Stock or Other Ownership: Celgene

Tommy Fu
Employment: Celgene
Stock or Other Ownership: Celgene

Bertrand Coiffier
Consulting or Advisory Role: Celgene, Celltrion, Erytech Pharma, MorphoSys, Mundipharma, Novartis, Pfizer
Research Funding: Celgene (Inst)
Expert Testimony: Gilead Sciences
Other Relationship: Celltrion, Celgene

Christopher R. Flowers
Consulting or Advisory Role: OptumRx, Seattle Genetics, Bayer HealthCare Pharmaceuticals, Gilead Sciences, Spectrum Pharmaceuticals, AbbVie, AstraZeneca, BeiGene, Denovo Biopharma, Karyopharm Therapeutics, Pharmacyclics/Janssen
Research Funding: Acerta Pharma (Inst), Infinity Pharmaceuticals (Inst), Onyx Pharmaceuticals (Inst), Janssen Oncology (Inst), Gilead Sciences (Inst), Celgene (Inst), TG Therapeutics (Inst), Genentech/Roche (Inst), Pharmacyclics (Inst), AbbVie (Inst), Immune Design (Inst), BeiGene (Inst)
Travel, Accommodations, Expenses: Celgene, Genentech/Roche
Acknowledgment

We thank and dedicate this manuscript to Dr Daniel Sargent, who was a principal leader in integrating his expertise in biostatistics and oncology with vision and foresight for establishing the importance of identifying early surrogate markers of prolonged survival outcomes. Dan was vital in bringing global investigators together to build data-sharing programs with large metadatabases across academic and industry teams for purposes of answering advanced research questions across multiple diverse studies, and with the continuation of these studies, his legacy continues to live on. We thank all the patients, families, and caregivers who participated in each of these studies, as well as all of the investigators, contributors, and study groups included in the analysis and for providing study data. We also thank all the SEAL group coinvestigators and study groups who contributed data for this analysis and the Mayo Clinic. Editorial support was provided by Julie Kern and Monica Nicosia with Bio Connections and funded by Celgene.

Appendix

The following investigators, contributors, and study groups were included in the analysis: Francesco Merli and the Intergruppo Italiano Linfomi (ANZINTER), Thomas M. Habermann and the US Intergroup (ECOG 4494), Groupe d’Etuve des Lymphomes de l’Adulte (GELA; now LYSARC), Amgen, and Roche, along with Christian Récher (LNH03-2B), Richard Delarue (LNH03-6B), Corrinne Haion (LNH 98-3), Bertrand Coiffier and Pierre Feugier (LNH98-5), John F. Seymour and F. Hoffman-La Roche (MAIN), Norbert Schmitz and the German High-Grade Lymphoma Study Group (DSHNHL 2002-1), Michael Pfreundschuh and the MabThera International Trial Group (MlnT) and Roche, Ulrich Jaeger and the Arbeitsgemeinschaft Medikamentöse Tumortherapie (AGMT; NHL13), Raoul Herbrecht and Cell Therapeutics (PIX203), Michael Pfreundschuh and Deutsche Krebshilfe (RICOVER-60), and David Cunningham and Cancer Research UK (UCL).


**Fig A1.** Distribution of (A) progression-free survival in overall population pooling of all patients across studies and (B) overall survival in overall population pooling of all patients across studies at a median follow-up time of 52 months.