Lymphoma

Association of progression-free or event-free survival with overall survival in diffuse large B-cell lymphoma after immunochemotherapy: a systematic review

Jie Zhu1,2 · Yong Yang1 · Jin Tao3 · Shu-Lian Wang1 · Bo Chen1 · Jian-Rong Dai1 · Chen Hu4 · Shu-Nan Qi1 · Ye-Xiong Li1

Received: 17 April 2020 / Accepted: 1 July 2020 / Published online: 10 July 2020
© The Author(s) 2020. This article is published with open access

Abstract
To investigate progression-free survival (PFS) and event-free survival (EFS) as early efficacy endpoints in diffuse large B-cell lymphoma (DLBCL), this systematic review included phase III randomized controlled trials (RCTs), phase II trials, and retrospective studies in newly diagnosed DLBCL receiving rituximab-containing chemotherapy through databases search up to 2019. Quality control was performed, where studies with high risk of bias were excluded. Prediction models were first established using the RCTs, and then externally validated in the phase II and retrospective populations. Trial-level surrogacy analysis was conducted by correlating the logarithmic (log) hazard ratio (HR) for PFS or EFS and log HR for OS. Correlation analysis at treatment arm-level was performed between 1-, 2-, 3-, and 5-year PFS or EFS rates and 5-year OS. The correlation was evaluated using the Pearson correlation coefficient $r$ in weighted linear regression, with weight equal to patient size. Sensitivity analyses were performed to assess the consistency of predictive model by leaving one subgroup of trials out at a time. Twenty-six phase III RCTs, 4 phase II trials and 47 retrospective studies were included. In trial-level surrogacy, PFS ($r$, 0.772; 95% confidence interval [CI], 0.471–0.913) or EFS ($r$, 0.838; 95% CI, 0.625–0.938) were associated with OS. For rituximab immunochemotherapy treatment arms in RCTs, there was a linear correlation between 1 and 5-year PFS ($r$, 0.813–0.873) or EFS ($r$, 0.853–0.931) and 5-year OS. Sensitivity analysis demonstrated reasonable overall consistency. The correlation between PFS and OS was externally validated using independent phase II, and retrospective data ($r$, 0.795–0.897). We recommend PFS and EFS as earlier efficacy endpoints in patients with DLBCL primarily treated with rituximab-containing immunochemotherapy.

Introduction
Diffuse large B-cell lymphoma (DLBCL) is the most common aggressive lymphoma subtype. Immunochemotherapy, mostly with rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP), has
become the standard treatment over the past decade [1–4]. However, 15–40% of patients are refractory to initial immunochemotherapy, or relapse after complete response (CR). Such patients have poor outcomes, mainly depending on the risk group [5]. There is an urgent need to find more effective agents or regimens for high-risk patients in the immunochemotherapy era.

Overall survival (OS) is the gold-standard treatment endpoint in randomized controlled trials (RCTs). However, OS as the primary endpoint requires a large sample size and long follow-up time to observe the survival benefit, leading to high clinical development costs and delays in introducing novel drugs. When used as the primary endpoints in clinical trials, early efficacy endpoints such as progression-free survival (PFS) and event-free survival (EFS) may require a smaller sample size and shorter evaluation time than OS, and have been established in some malignancies [6–8]. Trial- and individual-level studies have demonstrated that 24-month PFS and EFS may be considered the early efficacy endpoints for OS in DLBCL [9–12]. However, these studies may not be comprehensive because they only included available 13 RCTs willing to disclose individual patient data and were based on a subset of all potentially eligible trials [1–4, 12–21]. The association of PFS or EFS with OS has not been specifically addressed at trial- or treatment arm-level in RCTs on patients treated with immunotherapy; furthermore, its association and predictive value have not been externally validated. We investigated PFS and EFS as efficacy endpoints in DLBCL in the rituximab era through literature-based analysis at both trial- and treatment arm-level. The correlation between PFS and OS was validated in independent cohort studies to confirm its significant role in guiding clinical practice.

Methods

Literature search and study selection

Inclusion and exclusion criteria

This study was exempted from review by the institutional review board because it used existing data and enrolled no human subjects. The eligibility criteria included phase III RCTs, phase II trials, and retrospective studies investigating the long-term survival of DLBCL patients who received first-line rituximab-containing immunochemotherapy. Studies were excluded if they met any of the following conditions: phase I trial; transformed or relapsed/refractory DLBCL; inadequate survival data; serology-positive for HIV, hepatitis B/C virus, or Epstein–Barr virus; sample size of <100 patients per arm; or patients with DLBCL consisting of <80% of the whole-sample size.

Literature search

Studies published before 31 December 2019, were included via a systematic literature search of MEDLINE, Embase, and PubMed using the keyword “DLBCL AND rituximab” and with the restriction to RCT, phase II trial, and retrospective study. Formal publications and meeting abstracts were included. Two authors (J.Z. and J.T.) conducted the literature search independently, and reviewed the results with a third author (S.N.Q.). When disagreement in study inclusion was met, J.Z., J.T and S.N.Q. carefully reviewed the potential eligible study again. Disagreements about study inclusion were resolved by consensus.

RCT inclusion and quality control

All potentially eligible RCTs were assessed for risk of bias in seven domains (random sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective reporting, and other bias) using the Cochrane Collaboration tool. All information available in the assessment was acquired from formal publications, meeting abstracts, trial registry information on ClinicalTrials.gov (www.clinicaltrials.gov), and e-mail contact with trial designers. RCTs with high risk of bias in any domain were excluded.

A total of 109 abstracts were reviewed. After excluding 43 ineligible records, the full texts of 66 records were reviewed. Thirty-nine unqualified records were excluded, and 27 RCTs were included in the quality assessment (Figs. 1a and 2; Supplemental Table 1) [1–4, 13–19, 22–37]. Seven trials were rated with unclear risk of selection bias because of the lack of comprehensive reporting on the randomization process. The LNH03-1B trial was excluded from the study because of the high risk of bias related to its premature close and a sample size far below statistical requirements (Fig. 2; Supplemental Table 1) [37]. Eventually, 26 qualified RCTs were included for trial- and treatment arm-level analyses (Table 1) [1–4, 13–19, 22–36]. According to the purposes of each trial, 26 RCTs were classified into 5 subgroups: (1) four RCTs (15%) compared R-CHOP (like) with CHOP (like) [1–4]; (2) ten (38%) RCTs compared R-CHOP (like) with rituximab+intensified/desescalated chemotherapy [13, 15–17, 22, 23, 25–27, 31]; (3) nine (35%) investigated maintenance or consolidation therapy [3, 14, 18, 21, 24, 28–30, 32]; (4) three (12%) focused on R-CHOP+novel targeted therapy [19, 34, 35]; (5) two (8%) investigated the novel use of anti-CD20 monoclonal antibody [33, 36]. Of note, the two-stage randomized trial ECOG4494/CALGB9739 [3] and the 2×2 factorial randomized trial DCL04 [23] were classified into 2 subgroups according to the respective research questions. These 26 RCTs included a total of 16,340 patients (median sample
size, 623), with a median follow-up time of 2–10 years. The most common primary endpoints in these RCTs were EFS \((n = 12, 46\%)\) and PFS \((n = 7, 27\%)\), followed by disease free- or failure-free survival \((n = 5, 20\%)\), OS \((n = 1, 4\%)\), and CR \((n = 1, 4\%)\). The majority of RCTs \((n = 20, 77\%)\) used 2 or 3 years as the time point of the primary endpoint.

Phase II trial and retrospective study inclusion and quality control

To validate the RCT findings, we analyzed the relationship between PFS and OS using phase II and retrospective data. For single-arm phase II trials and retrospective cohort studies, quality was assessed, with a maximum 9-star score, using the Newcastle–Ottawa scale (NOS) in terms of selection, comparability, and outcome [38]. Studies with low to moderate risk of bias \((\geq 6\text{ stars})\) were included in the statistical analysis. For the LNH2007-3B randomized phase II trial [39], the risk of selection bias was assessed using the Cochrane Collaboration tool. A total of 1129 abstracts were reviewed. After excluding 865 unqualified records, the full texts of 264 records were reviewed. We excluded 203 ineligible studies, and included 61 studies in the quality assessment (Supplemental Table 2). After excluding 10 studies with high risk of bias, a total of 47 retrospective studies and 4 phase II trials with 67 rituximab immunochemotherapy treatment arms were included in the external validation (Fig. 1b) [39–89]. The average NOS score was 6.9 stars. A total of 14,936 patients were included, with each arm containing 100–1322 patients (median, 177). The median follow-up time was 1.2–7.2 years (Table 2).

Statistical methods

Endpoint definition

In the RCTs [1–4, 13–19, 22–36, 39], OS was defined as the time from randomization to death from any cause. EFS was defined heterogeneously, but generally from randomization to any treatment failure, including disease progression, death, and treatment discontinuity for any reason (e.g., adverse effects or withdrawal). PFS was generally measured from the time of randomization to disease progression, relapse, or death from any cause (Supplemental Table 3). In the retrospective studies [43–89], OS was generally defined as the time from diagnosis or treatment to death from any cause, and PFS from diagnosis or treatment to disease progression, relapse, or death from any cause (Supplemental Table 4).

Data extraction

In the RCTs, patient characteristics, sample size, follow-up period, primary endpoint, standard and treatment arms, hazard ratio (HR), absolute EFS/PFS rates (year 1, 2, 3, 5), and 5-year OS were extracted (Table 1). For a repeatedly reported RCT, we included the most recent result with the longest follow-up time. All results of the standard and treatment arms were based on the intention-to-treat population. For the phase II trials and retrospective studies, patient characteristics, sample size, median follow-up time, treatment, absolute PFS rates (year 1, 2, 3, 5) and 5-year OS were extracted (Table 2). As described previously [90], the HR or survival rates at the different time points was extracted from the full text (labeled “*”) or the Kaplan–Meier survival curve using Engauge Digitizer software.

Correlation evaluation

The correlation analyses of the RCTs, weighted by trial size, were performed at both trial- and rituximab immunochemotherapy arm-level, without inclusion of treatment arms using conventional CHOP (like) regimen in arm-level analysis. At trial-level, the correlation of \( \log \text{HR (PFS)} \) or \( \log \text{HR (EFS)} \) with \( \log \text{HR (OS)} \) was estimated using the Pearson correlation coefficient \( r \) in weighted linear regression, with weight equal to trial sample size. At rituximab
immunochemotherapy arm-level, the linear correlation between the 1-, 2-, 3-, and 5-year PFS or EFS rates and 5-year OS rate was also evaluated by the correlation coefficient $r$, with weight depending on the sample size of each treatment arm. A strong association was indicated when the value of $r$ was close to 1, and the 95% confidence intervals (CIs) of $r$ were obtained using the bootstrap method with 1000 replications.

**Sensitivity analysis**

Phase III RCTs were classified into five subgroups according to study purposes. To assess the consistency and robustness of the developed predictive model across different settings, sensitivity analyses were performed by leaving each subgroup of trials out at a time. The correlation coefficient $r$ and its 95% CI in trial-level and treatment arm-level correlation were reported similarly.

**External validation of RCT prediction model in phase II trials and retrospective studies**

We validated our finding by applying the predictive linear regression models to the phase II and retrospective studies with adequate survival data. The predicted 5-year OS rate was calculated from the actual 1–5-year PFS rates in the phase II or retrospective studies using the established linear regression model from the RCTs. For example, the equation “5-year OS = $\alpha \times 1-, 2-, 3-, or 5$-year PFS + $\beta$” was derived from the RCTs. Using the reported 1–5-year PFS rate derived from the phase II and retrospective studies, we used these models to generate the predicted 5-year OS rates. The actual and predicted 5-year OS rates were plotted in scatter plots. Statistical analysis was performed in SPSS (version 21.0, IBM Inc.); data visualization was performed using the ggplot2 package in R software (version 3.3.2, R Foundation for Statistical Computing).

**Data sharing statement**

For original data, please contact yexiong12@163.com.

**Results**

**Trial-level correlation between treatment effects of PFS or EFS on OS in RCTs**

Of 26 RCTs (Table 1), 20 (77%), 1 (4%), and 1 (4%) reported one, two, and three pairs of PFS HR and OS HR, respectively. A significant correlation was observed after

---

**Fig. 2 Summary of risk of bias in RCTs.** “+” (green), “?” (yellow), and “−” (red) represent low, unclear, and high risk of bias, respectively. RCTs randomized controlled trials.
| Trial                  | Inclusion Criteria                                      | Primary endpoint | Median PU year | No. | Treatment   | PFS, % | EFS, %   | OS, % |
|-----------------------|--------------------------------------------------------|------------------|----------------|-----|-------------|--------|----------|-------|
|                       |                                                        |                  |                |     | HR          | 1-y   | 2-y      | 3-y   | 5-y   |
|                       |                                                        |                  |                |     |             |        |          |       |       |
|                       |                                                        |                  |                |     | HR          | 1-y   | 2-y      | 3-y   | 5-y   |
|                       |                                                        |                  |                |     |             |        |          |       |       |
| R-CHOP (like) vs. CHOP (like) (n = 4) |                                                        |                  |                |     |             |        |          |       |       |
| LNH-98.5 [1]          | Age 60–80; stage II–IV; PS ≤ 2                        | 3-y EFS          | 3.5            | 3.0 | R-CHOP     | 0.61P | 67.3     | 53     | 47*    | 0.69P | 58* |
|                       |                                                       |                  |                |     | CHOP       | 0.67R | 67.3     | 53     | 47*    | 0.69P | 58* |
| CML [2]               | Age 18–60; allPI ≤ 1; stage II–IV or bulky stage I    | 3-y EFS          | 3.5            | 3.0 | R-CHOP-like| 0.48P | 67.3     | 53     | 47*    | 0.69P | 58* |
|                       |                                                       |                  |                |     | CHOP-like  | 0.67R | 67.3     | 53     | 47*    | 0.69P | 58* |
| ECOCOG4494/ CALGB7939 [3] | 1st randomization: age ≥ 60; all stage; PS ≤ 3  | 3-y FFS          | 3.5            | 3.0 | R-CHOP     | NA    | 63.2     | 46*    | 47.7   |       |     |
|                       |                                                       |                  |                |     | CHOP       | NA    | 63.2     | 46*    | 47.7   |       |     |
| RICOVER-60 [4]        | Age 61–80; all stages                                  | 3-y EFS          | 3.5            | 2.9 | R-CHOP     | NA    | 63.2     | 46*    | 47.7   |       |     |
|                       |                                                       |                  |                |     | CHOP       | NA    | 63.2     | 46*    | 47.7   |       |     |
| R-CHOP (like) vs. R+ intensified/ide-escalated chemotherapy (n = 10) |                                                        |                  |                |     |             |        |          |       |       |
| LNH03-2B [13]         | Age 18–59; all stages; allPI 1                        | 2-y EFS          | 3.5            | 3.0 | R-CHOP     | 0.61P | 67.3     | 53     | 47*    | 0.69P | 58* |
|                       |                                                       |                  |                |     | CHOP       | 0.67R | 67.3     | 53     | 47*    | 0.69P | 58* |
| ANZINTER3 [15]        | Age > 65; stage II–IV; PS ≤ 3 in CGA                  | 2-y EFS          | 3.5            | 3.0 | R-CHOP     | 0.61P | 67.3     | 53     | 47*    | 0.69P | 58* |
|                       |                                                       |                  |                |     | CHOP       | 0.67R | 67.3     | 53     | 47*    | 0.69P | 58* |
| LNH03-6B [16]         | Age 60–80; allPI ≥ 1                                  | 2-y EFS          | 3.5            | 3.0 | R-CHOP     | 0.61P | 67.3     | 53     | 47*    | 0.69P | 58* |
|                       |                                                       |                  |                |     | CHOP       | 0.67R | 67.3     | 53     | 47*    | 0.69P | 58* |
| NCT1793844 [22]       | Age ≥ 18; all stages; PS ≤ 3                          | 3-y DFS          | 3.5            | 3.0 | R-CHOP     | 0.61P | 67.3     | 53     | 47*    | 0.69P | 58* |
|                       |                                                       |                  |                |     | CHOP       | 0.67R | 67.3     | 53     | 47*    | 0.69P | 58* |
| UK NCRI [17]          | Age ≥ 18; stage III–IV; bulky IA; PS ≤ 2             | 2-y OS           | 3.5            | 3.0 | R-CHOP     | 0.61P | 67.3     | 53     | 47*    | 0.69P | 58* |
|                       |                                                       |                  |                |     | CHOP       | 0.67R | 67.3     | 53     | 47*    | 0.69P | 58* |
| DLCL04 [23]           | Age 18–65; DLBCL or FL 3b; stage II–IV; PS ≤ 3; high or intermediate-risk | 2-y FFS          | 3.5            | 3.0 | R-MegaCHOP | NA    | 63.2     | 46*    | 47.7   |       |     |
|                       |                                                       |                  |                |     | 6 CHOP     | NA    | 63.2     | 46*    | 47.7   |       |     |
| Alliance/CALGB 50303 [25] | Age ≥ 18; stage I (PMBCL) or II–IV; PS ≤ 2          | PFS              | 5.5            | 5.5 | DA-EPOCH-R | 0.93N | 88.9     | 75.8   | 68*    | 1.05N | 77.5 |
|                       |                                                       |                  |                |     | R-CHOP     | 0.93N | 88.9     | 75.8   | 68*    | 1.05N | 77.5 |
| FLYER [26]            | Age 18–60; stage I–II; PS ≤ 1; tumor < 7.5 cm        | 3-y PFS          | 5.5            | 5.5 | R-CHOP     | 0.93N | 88.9     | 75.8   | 68*    | 1.05N | 77.5 |
|                       |                                                       |                  |                |     | CHOP       | 0.93N | 88.9     | 75.8   | 68*    | 1.05N | 77.5 |
| PETAL [31]            | Age 18–80; B- or T-cell; PET (–)                      | 2-y EFS          | 4.5            | 4.5 | R-CHOP     | 0.93N | 88.9     | 75.8   | 68*    | 1.05N | 77.5 |
|                       |                                                       |                  |                |     | CHOP       | 0.93N | 88.9     | 75.8   | 68*    | 1.05N | 77.5 |
| NHL-001 [27]          | Age 61–80; DLBCL or FL 3b; all stages; PS ≤ 2        | 2-y PFS          | 3.5            | 3.5 | R-CEOP90   | 0.44N | 88.9     | 75.8   | 68*    | 1.05N | 77.5 |
|                       |                                                       |                  |                |     | R-CEOP70   | 0.44N | 88.9     | 75.8   | 68*    | 1.05N | 77.5 |
|                       |                                                       |                  |                |     | R-CEOP50   | 0.44N | 88.9     | 75.8   | 68*    | 1.05N | 77.5 |
|                       |                                                       |                  |                |     | R-CHOP     | 0.93N | 88.9     | 75.8   | 68*    | 1.05N | 77.5 |
|                       |                                                       |                  |                |     | CHOP       | 0.93N | 88.9     | 75.8   | 68*    | 1.05N | 77.5 |
| R-CHOP (like) chemotherapy followed by maintenance consolidation therapy (n = 9) |                                                        |                  |                |     |             |        |          |       |       |
| AGMT-NHL13 [18]       | Age ≥ 18; all stages; PS ≤ 3-y EFS 2; CR/CRu         | 3-y EFS          | 3.5            | 3.5 | R-CHOP-like|R 0.62P | 88.9     | 83.7   | 80*    | 0.81N | 90.7 |
|                       |                                                       |                  |                |     | maintenance|        |          |       |       |
| Trial                          | Inclusion Criteria                                                                 | Primary endpoint | Median FU, year | No. | Treatment                                      | Treatment codes                                      | PFS, % | EFS, % | OS, % | Treatment codes                                      | HR 1-y | 2-y | 3-y | 5-y | HR 1-y | 2-y | 3-y | 5-y | HR 1-y | 2-y | 3-y | 5-y | HR 5-y |
|-------------------------------|------------------------------------------------------------------------------------|------------------|----------------|-----|-----------------------------------------------|------------------------------------------------------|--------|--------|-------|-----------------------------------------------|---------|----|----|----|---------|----|----|----|---------|----|----|----|--------|
| ECOG4494/ CALGB9791 [3]       | 2nd randomization; age ≥ 60; all stage; PS ≤ 3; CR/PR                             | 2-y FFS          | 3.5            | 345 | R-CHOP-like                                    | R maintenance                                       | 87.4   | 84.1  | 79*   | 68.8                                          | 0.63*   |     |     |     | 88.8   | 81.5| 76.5* | 61.4 | 64.0    |     |     |     |     |
| PRELUDE [28]                 | Age ≥ 18; stage bulky II or III or IV; PS ≤ 2; high risk; CR/CR                      | 3-y DFS          | 3              | 493 | R-CHOP+enzastaurin maintenance                 | 0.9*N (DFS)                                         | 82.3*  | 78.1  | 72.7* | NA                                            | 71      | 61* | 56.7 | 52.7 | 1.04*   |     |     |     |     |
| REMARC [29]                  | Age 60–80; stage II–IV; aaIPI ≥ 1; PS ≤ 2; CR or PR                                | 2-y FFS          | 3.3 (PFS)      | 323 | R-CHOP+lenalolodimide maintenance              | 0.71* (P)                                           | 87.8   | 80*   | 74.4  | 61.4                                          | 86.7    |     |     |     | 76.4    |     |     |     |     |
| PILLAR-2 [30]                | Age > 18; stage II (tumor > 10 cm) or III or IV; CR                               | 2-y DFS          | 4.2            | 372 | R-Chemo+everolimus maintenance                | 0.92* (DFS)                                         | 81.6   | 75*   | 67    | 49.3                                          | NA      |     |     |     | 77.7    |     |     |     |     |
| DSHNHL2002-1 [14]            | Age < 61; stage III–IV; aaIPI 2–3; high risk                                     | 3-y EFS          | 3              | 370 | R-Chemo                                       |                                                     | 80.6   | 77*   | 70.7  | 60.8                                          | NA      |     |     |     | 72.5    |     |     |     |     |
| DLCLI4 [23]                  | Age 18–65; DLBCL or FL 3b; stage II–IV; aaIPI 2–3; PS ≥ 2; high or intermediate-high risk | 2-y FFS          | 6              | 199 | R-HDC+ASCT                                    |                                                      | 80.5   | 72*   | 70*   | 69.8                                          | 74.9    |     |     |     | 52.5    |     |     |     |     |
| NCT00355199 [24]             | Age 18–65; stage III–IV or bulky II; high risk                                   | 3-y EFS          | 5              | 113 | R-HDC+ASCT                                    |                                                      | 70.6   | 65*   | 59*   | 54.8                                          | 70.2    |     |     |     | 77*     |     |     |     |     |
| LYSA/GOELAMS [32]            | Age: 18–75; stage I–II; tumor < 7 cm                                             | 5-y EFS          | 5.3            | 165 | R-Chemo+ RT                                   |                                                      | 75.5   | 73.1  | 70.1  | 65.2                                          | 68.4    |     |     |     | 60.1    |     |     |     |     |
| R-CHOP+novel targeted drug (n = 3) |                                    |                  |                |     |                                              |                                                      |        |       |       |                                               |         |     |     |     |         |     |     |     |        |
| MAIN [19]                    | Age ≥ 18; all stages                                                               | 2               | PFS to safety  | 2   | R-CHOP                                        |                                                      | 75.1   | 68.3  | 60.8  | NA                                            | NA      |     |     |     | 1.03*   |     |     |     |     |
| REMoDL-B [34]                | Age ≥ 18; stage I (tumor > 10 cm) or II–IV; PS ≤ 2; with GEP                       | 2.5-y PFS        | 2.5            | 358 | R-RB-CHOP                                     |                                                      | 81.1   | 76.1  | 75.3  | 68.1                                          | 78.8    |     |     |     | 64.0    |     |     |     |     |
| PHOENIX [35]                 | Age ≥ 18; non-GCB; stage II–IV; R-IPI ≥ 1; PS ≤ 2                                 | EFS             | 2.9            | 419 | R-Chop +brutinib                              |                                                      | 81.3   | 73.3  | 69.6  | 69.6*                                         | 81.3    |     |     |     | 92*     |     |     |     |     |
| Anti-CD20 monoclonal antibody study (n = 2) |                                    |                  |                |     |                                              |                                                      |        |       |       |                                               |         |     |     |     |         |     |     |     |        |
| MabEase [36]                 | Age 18–80; tumor ≥ 1.5 cm; IPI 0 ≥ 7.5 cm or 1–5; PS ≤ 2                         | CR/CRu           | 2.9            | 381 | (R)SC+CHOP                                    |                                                      | 85.2   | 75*   | 70   | 61.8                                          | 77      |     |     |     | 35.8    |     |     |     |     |
| GOYA [33]                    | Age ≥ 18; tumor > 1.5 cm; PS ≤ 2; IPI 0 ≥ 7.5 cm or I (age < 60 or ≥ 62); LVIF ≥ 50% | 3-y PFS          | 2.4            | 706 | G-Chop                                        |                                                      | 81.3   | 73    | 69.6  | 69.6*                                         | 79.9    |     |     |     | 66.9*   |     |     |     |     |

The standard arm is labeled in bold.
analyzing 25 pairs of PFS HR and OS HR. Log HR (PFS) correlated with log HR (OS) \( (r = 0.772; \ 95\% \ CI, 0.471–0.913; \) Fig. 3a). Sensitivity analyses showed good consistency in most subgroups, except when leaving the subgroup R-CHOP (like) vs. CHOP (like) out \( (r = 0.61; \ 95\% \ CI, 0.075–0.863; \) Supplemental Fig. 1a). This result was expected. Among the 26 RCTs we studied, 4 trials \([1, 2, 4, 13]\) were shown statistically significant OS benefits, including 3 trials \([1, 2, 4]\) in the subgroup comparing R-CHOP (like) with CHOP (like). The exclusion of these positive trials at once naturally leads to a wider confidence interval.

Fourteen RCTs (54%) reported one pair of EFS HR and OS HR each (two treatment arms); three RCTs (12%) reported two pairs of EFS HR and OS HR each (four treatment arms). The analysis of 20 pairs of EFS HR and OS HR demonstrated that log HR (EFS) correlated with log HR (OS) \( (r = 0.838; \ 95\% \ CI, 0.625–0.938; \) Fig. 3b). Sensitivity analyses demonstrated good consistency in most subgroups, except when leaving the subgroup R-CHOP (like) vs. CHOP (like) out \( (r = 0.732; \ 95\% \ CI, 0.278–0.941) \) because of similar reasons as in PFS (Supplemental Fig. 1b). These results confirm that treatment gain in PFS or EFS can predict OS benefit at trial level with an acceptable consistency.

**Treatment arm-level correlation between PFS or EFS and OS in RCTs**

Forty-four rituximab immunochemotherapy arms from 26 RCTs reported 5-year OS. Thirty-five (80%) rituximab immunochemotherapy arms reported 1-year and 3-year PFS; 37 (84%) arms reported 2-year PFS and 33 (75%) arms reported 5-year PFS. The 1-year \( (r = 0.813; \ 95\% \ CI, 0.624–0.913; \) Fig. 4a), 2-year \( (r = 0.858; \ 95\% \ CI, 0.705–0.933; \) Fig. 4b), 3-year \( (r = 0.873; \ 95\% \ CI, 0.716–0.946; \) Fig. 4c), or 5-year PFS \( (r = 0.871; \ 95\% \ CI, 0.711–0.954; \) Fig. 4d) correlated linearly with 5-year OS. Generally speaking, sensitivity analyses continued to demonstrate robust consistency in terms of correlation r. When leaving out 10 trials from R-CHOP (like) with rituximab-intensified/de-escalated chemotherapy subgroup (Supplemental Fig. 1c–f), which account for nearly half of all treatment arms, the findings remain consistent with wider confidence intervals due to the reduced number of arms.

Twenty-seven rituximab immunochemotherapy arms (61%) reported 1-, 2-, 3- and 5-year EFS. Linear regression analysis revealed correlations between 1-year \( (r = 0.853; \ 95\% \ CI, 0.729–0.920; \) Fig. 5a), 2-year \( (r = 0.896; \ 95\% \ CI, 0.815–0.945; \) Fig. 5b), 3-year \( (r = 0.921; \ 95\% \ CI, 0.851–0.966; \) Fig. 5c), or 5-year EFS \( (r = 0.931; \ 95\% \ CI, 0.885–0.975; \) Fig. 5d) and 5-year OS. Sensitivity analysis indicated good consistency (Supplemental Fig. 1g–j). This
Table 2: Summary of phase II and retrospective studies used for predictive model validation.

| Study | NOS | Eligibility | Treatment | No. | FU (Years) | 1-y | 2-y | 3-y | 5-y |
|-------|-----|-------------|-----------|-----|------------|-----|-----|-----|-----|
| **Phase II trial (n = 4)** | | | | | | | | | |
| LNHH2007-3B [39] | NA | Age 18–59; aaIPI 2–3 | R-ACVBp | 109 | 3.8 | 83.7 | 80.3 | 76.6 | 75.2 | 84.8 |
| | | | R-CHOP | 102 | | 80.9 | 76.3 | 74.3 | 74.4 | 80.3 |
| DENSE-R-CHOP-14 [40] | 9 | Age 61–80 | R-CHOP+6 R | 124 | 4.3 | 81.7 | 73.3 | 67.0* | 55.7 | 62.3 |
| LNHH2003-3 [41] | 9 | Age 18–60; aaIPI 2 | R-ACVBp+ASCT | 157 | | 84.8 | 79.6 | 77.8 | 76.5 | 78.6 |
| Niitsu N, et al. [42] | 9 | Age 15–60; stage II–IV | R-CyclOBEAP | 101 | 3.5 | 95.2 | 79.7 | 76.3 | 76.0* | 85.0* |
| **Retrospective study (n = 47)** | | | | | | | | | |
| Go SI, et al. [43] | 7 | PNI ≥ 40 | R-CHOP | 159 | 5.8 | 82.3 | 73.9 | 70.7 | 65.9 | 69.8 |
| Lee J, et al. [44] | 7 | GCB | Non-GCB | 120 | 1.2 | 82.6 | 80.1 | 75.8 | 70.0* | 71.0* |
| | | | R-chemo | 177 | | 77.4 | 69.8 | 68.0 | 65.0* | 70.0* |
| Morrison VA, et al. [45] | 7 | All stages | R-chemo | 1322 | 1.9 | 79.6 | 68.3* | 63.1 | 47.7 | 67.4 |
| Yim SK, et al. [46] | 8 | PET/CT score 1–3 | R-CHOP | 171 | 4.7 | 87.6 | 83.3 | 77.9 | 72.6* | 78.1* |
| Chen Y, et al. [47] | 6 | BM PET/CT (–) | R-CHOP | 147 | 2.5 | 87.9 | 82.6 | 81.5* | 77.0 | 88.4 |
| Hosoda Y, et al. [48] | 7 | All stages | R-chemo | 182 | 3.7 | 74.2 | 70.5 | 66.0* | 52.1 | 66.6 |
| Kim SH, et al. [49] | 7 | AGR ≥ 1.22 | R-CHOP | 139 | 5.5 | 82.2 | 74.9 | 73.0 | 69.8 | 70.8 |
| Li LY, et al. [50] | 7 | BCL2 (+) | R-chemo | 145 | 1.9 | 48.6 | 38.5 | 36.2 | 30.0* | 45.2 |
| Li YW, et al. [51] | 7 | Uric acid < 6.4 mg/dL | R-CHOP or like | 114 | 1.8 | 90.8 | 86.4 | 83.5 | 82.8 | 83.0 |
| Matsumoto K, et al. [52] | 7 | All stages | R-CHOP | 185 | 4.6 | 85.4 | 80.5 | 76.1* | 72.0* | 80.1* |
| Sun FF, et al. [53] | 7 | ICPS 0 | R-CHOP | 202 | 2.6 | 93.2 | 86.9 | 86.5* | 81.8 | 91.8 |
| | | ICPS 1 | R-CHOP | 144 | | 87.9 | 83.8 | 82.3* | 78.2 | 87.0 |
| | | ICPS 3 | R-CHOP | 119 | | 65.6 | 60.0 | 54.5* | 49.1 | 58.8 |
| Go SI, et al. [54] | 7 | Sarcopenia-L3 | R-CHOP | 141 | 4.9 | 78.9 | 70.4 | 69.8 | 64.9* | 67.8* |
| Kanemasa Y, et al. [55] | 7 | B2MG ≥ 3.2 mg/L | R-CHOP or like | 101 | 3.1 | 62.6 | 46.8 | 45.3* | 35.7 | 41.2 |
| Li J, et al. [56] | 7 | AA genotype of EP300 SNP rs20551 | R-CHOP | 192 | 5.3 | 80.0 | 71.5 | 69.2 | 68.6* | 77.0* |
| Liu YL, et al. [57] | 7 | TP53 Arg72 | R-CHOP | 238 | 4.7 | 77.1 | 67.6 | 64.0 | 63.5* | 74.9* |
| Park YH, et al. [58] | 7 | High ALI | R-CHOP | 130 | 4.6 | 91.1 | 85.1 | 78.8 | 77.3* | 80.2* |
| Song MK, et al. [59] | 7 | No tumor necrosis | R-CHOP | 387 | 4.1 | 86.0 | 75.5 | 72.3 | 68.3* | 74.3* |
| Tsuyama N, et al. [60] | 6 | MYC (–), BCL2 (–) | R-CHOP | 179 | NA | 85.2 | 78.3 | 75.5 | 69.0 | 81.2 |
| Alinari L, et al. [61] | 7 | CD5+ | R-chemo | 102 | 3.3 | 64.0 | 43.2 | 40.0* | 40.0* | 60.0* |
| Prochazka KT, et al. [62] | 6 | Uric acid ≥ 6.8 mg/dL | R-chemo | 130 | NA | 71.2 | 58.3 | 54.8 | 44.0* | 50.4* |
| | | Uric acid < 6.8 mg/dL | R-chemo | 399 | | 79.7 | 70.6 | 66.9 | 59.6* | 66.2* |
| Seo S, et al. [63] | 7 | B2M ≥ 2.5 mg/L | R-CHOP | 290 | 4 | 59.0 | 49.2 | 44.6 | 41.0* | 49.2* |
| | | B2M < 2.5 mg/L | R-CHOP | 543 | | 88.8 | 85.2 | 83.0 | 76.1* | 83.8* |
| Dabaja BS, et al. [64] | 7 | All stages | R-CHOP+RT | 293 | 4.5 | 96.4 | 90.8 | 88.5 | 83* | 91.0* |
| | | All stages | R-CHOP | 548 | | 93.9 | 87.3 | 83.3 | 76.0* | 83.0* |
| El-Galaly TC, et al. [65] | 6 | IPI 0–1 | R-CHOP like | 138 | 2.4 | 93.4 | 89.2 | 89.1 | 86.8 | 90.7 |
| | | IPI 2 | R-CHOP like | 116 | | 86.1 | 80.9 | 73.0 | 60.8 | 70.2 |
| Gong QX, et al. [66] | 7 | CD30 (–) | R-CHOP | 112 | 2.9 | 69.9 | 55.5 | 52.2 | 48.2 | 60.6 |
| Kumar A, et al. [67] | 7 | Stage I/II | R-CHOP±RT | 261 | 4.7 | NA | NA | NA | 82.0* | 93.2* |
| Melchardt T, et al. [68] | 6 | NCCN-IPI 2-3 | R-CHOP or like | 199 | 4.3 | 85.3 | 77.4 | 73.9* | 68.8* | 77.3* |
| | | NCCN-IPI 4-5 | R-CHOP or like | 189 | | 78.5 | 66.6 | 63.5* | 52.2* | 56.4* |
finding indicates that improvements in 1–3-year PFS or EFS are associated with higher 5-year OS.

**External validation of association of PFS with OS in Phase II and retrospective studies**

Sixty-seven treatment arms from the phase II and retrospective studies were used for external validation. As EFS was not available in the retrospective studies, only PFS prediction models could be evaluated. Using the PFS predictive models from the RCTs (Fig. 4), we calculated the predicted 5-year OS rate for each retrospective study using the actual 1-, 2-, 3-, or 5-year PFS rate (Table 2). The simple regression line between the actual and predicted 5-year OS approached the diagonal line, indicating that the predicted OS was approximated to the actual OS. The

**Table 2 (continued)**

| Study | NOS (Stars) | Eligibility | Treatment | No. | FU (Years) | 1-y | 2-y | 3-y | 5-y |
|-------|-------------|-------------|-----------|-----|------------|-----|-----|-----|-----|
| Nakajima Y, et al. [69] | 7 | Stage II/II, supradiaphragm | R-CHOP | 109 | 4.3 | 90.9 | 88.8 | 86.6 | 86.4 | 92.2 |
| Dabaja BS, et al. [70] | 7 | PET/CT (−) | R-chemo | 239 | 3 | 89.0 | 84.3 | 81.0 | 78.0* | 82.0* |
| Mian M, et al. [71] | 6 | All stages | R-CHOP | 218 | 3.3 | 70.7 | 62.8 | 55.3 | 44.9 | 71.5 |
| All stages | R-COMP | 146 | 1.5 | 74.7 | 60.0 | 56.3 | 50.4 | 61.6 |
| Castillo JJ, et al. [72] | 8 | Asian patients | R-CHOP | 455 | 3 | 83.4 | 72.7 | 65.0 | 60.0* | 66.0* |
| Western patients | R-CHOP | 257 | | 80.5 | 69.2 | 65.0 | 55.0* | 64.0* |
| Hashimoto Y, et al. [73] | 6 | sIL-2Rα < 1000 U/mL | R-CHOP | 101 | 2.2 | 90.0 | 83.3 | 83.4 | 82.0 | 84.0 |
| Kojima M, et al. [74] | 7 | All stages | R-chemo | 100 | 4.2 | 78.0 | 66.3 | 62.0* | 61.4 | 66.0* |
| Lu HJ, et al. [75] | 6 | Stage I-III | R-CHOP | 232 | 3.3 | 60.9 | 58.4 | 57.8 | 57.3* | 69.8* |
| Ozbak M, et al. [76] | 6 | All stages | R-CHOP | 258 | 3.3 | 70.7 | 62.8 | 55.3 | 44.9 | 71.5 |
| Mian M, et al. [71] | 6 | All stages | R-COMP | 146 | 1.5 | 74.7 | 60.0 | 56.3 | 50.4 | 61.6 |
| Castillo JJ, et al. [79] | 6 | GC type | R-CHOP | 379 | NA | 86.2 | 79.6 | 75.9 | 67.4 | 68.5 |
| Non-GC type | R-CHOP | 333 | | 81.3 | 73.5 | 70.0 | 63.6 | 64.0 |
| Huang HH, et al. [80] | 6 | Age: 15–60 y; IPI ≥ 2 | R-CHOP | 112 | 5 | 94.7 | 69.7 | 50.3 | 40.9* | 56.7* |
| Li ZM, et al. [81] | 6 | LMR > 2.6 | R-CHOP | 280 | NA | 91.3 | 85.8 | 83.6 | 79.4* | 84.0 |
| LMR ≤ 2.6 | R-CHOP | 158 | | 76.6 | 66.6 | 60.8 | 54.2 | 64.6 |
| Li XY, et al. [82] | 6 | Age > 60 | R-chemo | 159 | NA | 59.4 | 49.0 | 45.4 | 41.6 | 42.0 |
| Tomita N, et al. [84] | 7 | Revised IPI 1-2 | R-CHOP | 201 | 3.6 | 88.5 | 82.6 | 81.7 | 79.0* | 89.0* |
| Revised IPI 3-5 | R-CHOP | 117 | | 73.6 | 60.9 | 58.1 | 56.0* | 63.0* |
| Sehn LH, et al. [85] | 7 | no BM involvement | R-CHOP | 670 | 3.4 | 82.4 | 76.4 | 73.0* | 68.5 | 73.5 |
| Bari A, et al. [86] | 7 | All stages | R-chemo | 271 | 3.4 | 78.3 | 68.0 | 65.5 | 53.2 | 60.0 |
| Enhishi D, et al. [87] | 7 | All stages | R-CHOP | 221 | 3.4 | 80.5 | 76.3 | 73.0 | 72.9 | 77.8 |
| Phan J, et al. [88] | 7 | All stages | R-CHOP+RT | 142 | 3 | 99.6 | 95.5 | 94.9 | 82.0* | 91.0* |
| All stages | R-CHOP | 327 | | 93.5 | 85.1 | 77.4 | 59.0* | 68.0* |
| Scandurra M, et al. [89] | 6 | Without del (8p23·1) | R-CHOP | 144 | 1.9 | 87.7 | 75.6 | 64.8 | 60.0 | 83.2 |

*Represents data directly reported in the full text.

Chemotherapy regimens: R, rituximab; R-ACVBP, rituximab, doxorubicin, cyclophosphamide, vindesine, bleomycin, and prednisone; R-chemo, rituximab-based chemotherapy; R-CHOP, rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone; R-COMP, rituximab, cyclophosphamide, non-pegylated liposomal doxorubicin, vincristine, and prednisone; R-CycloBEAP, rituximab, cyclophosphamide, vincristine, bleomycin, etoposide, doxorubicin, and prednisolone. aIPI age-adjusted International Prognostic Index, AGR albumin globulin ratio, ALI advanced lung cancer inflammation index, Arg72 arginine at codon 72, ASCT autologous stem cell transplantation, B2M beta-2 microglobulin, BM bone marrow, DM diabetes mellitus, FU follow-up, GC germinal center, GCB germinal center B-cell, ICPS inflammation-based cumulative prognostic score, IPI International Prognostic Index, LMR lymphocyte-to-monocyte ratio, NA not available, NCCN-IPI National Comprehensive Cancer Network–IPI, No. number of patients, NOS Newcastle–Ottawa scale, OS overall survival, PET/CT positron emission tomography/computed tomography, PFS progression-free survival, PNI prognostic nutritional index, RT radiotherapy, sIL-2Rα soluble interleukin-2 receptor-α, SNP single-nucleotide polymorphism.
predicted 5-year OS rate correlated significantly with the actual 5-year OS rate, with the correlation coefficient $r$ ranging from 0.795 to 0.897 (Fig. 6a–d). This finding validates the premise that PFS is predictive of OS.

**Discussion**

This is a large-scale, comprehensive study combining data from high-quality phase III RCTs, phase II trials, and retrospective studies to assess the association between the early efficacy endpoints of PFS or EFS with OS in patients with DLBCL primarily treated with immunochemotherapy. Consistent with previous findings [9–12], analyses of the 26 qualified RCTs showed that improved PFS or EFS correlated with OS benefit at trial level. There was a linear correlation between 1–5-year PFS or EFS and 5-year OS rates at treatment arm level. The comprehensive sensitivity analyses indicated an acceptable overall consistency of the developed predictive model across settings. The external validation showed good calibration between the actual and predicted 5-year OS rates based on the 1–5-year PFS rates in the phase II and retrospective studies. These findings provide new evidence supporting the clinical use of PFS and EFS as early efficacy endpoints for evaluating treatment benefit and accelerating approval for superior treatments.

Previous studies, primarily using 13 RCTs conducted before 2015, concluded that the early efficacy endpoints of EFS or PFS are strongly related to OS at both individual and trial level [9–11]. The survival of DLBCL patients who...
achieved PFS or EFS at 24 months is almost equal to that of the age- and sex-matched general population [9–12]. Therefore, 2-year EFS or PFS are accepted as early efficacy endpoints. Although the use of individual patient data allows better characterization of important covariates that affect survival, it restricts the analysis to a limited number of patients in each treatment arm. The solid blue line indicates the fitted weighted linear regression line; the light green zone represents its 95% CI; \( r \) indicates the correlation coefficient. EFS event-free survival; OS overall survival; RCTs randomized controlled trials; CI confidence interval.

Fig. 6 External validation of association of PFS with OS after Rituximab immunochemotherapy. Using PFS linear regression models (as shown in Fig. 4), the predicted 5-year OS, as calculated according to the actual 1-, 2-, 3-, and 5-year PFS from the phase II trials and retrospective data (Table 2), is plotted against the actual 5-year OS. The predicted OS approximates to the actual OS, as indicated by approaching the diagonal line, i.e., the line of identity; \( r \) indicates the correlation coefficient. PFS progression-free survival; OS overall survival.

Fig. 5 Rituximab Immunochemotherapy Arm-level Correlation Between EFS and OS in RCTs. The rituximab immunochemotherapy arm-level associations between a 1-year EFS and 5-year OS, b 2-year EFS and 5-year OS, c 3-year EFS and 5-year OS, and d 5-year EFS and 5-year OS. Circle size is proportional to the number of patients in each treatment arm. The solid blue line indicates the fitted weighted linear regression line; the light green zone represents its 95% CI; \( r \) indicates the correlation coefficient. EFS event-free survival; OS overall survival; RCTs randomized controlled trials; CI confidence interval.
RCTs, and the analysis is not easily replicated by independent researchers. In most recently published trials and in clinical practice, there are multiple effective agents not only as initial treatment but also in second-line or salvage settings. Any validation of an early efficacy endpoint is relevant only within the context in which the validation occurred. These factors prompted re-examination and external validation of the correlation between PFS or EFS at the given time points with OS. The present literature-based analysis relied on data from RCTs, phase II trials, and retrospective studies to assess the validity of the early efficacy endpoints, and represents a critical step toward understanding the impact of immunochemotherapy on PFS or EFS and OS in DLBCL. With strict inclusion criteria and quality control, we included large-scale, qualified RCTs for trial-level surrogacy analysis, and phase II trials and retrospective studies for external validation. The correlation between PFS or EFS with OS was well established for DLBCL at both the trial and treatment arm level from the RCTs. Furthermore, the correlation between 1–5-year PFS and OS was externally validated by analyzing the phase II and retrospective data. Consistent with previous studies [9–12], these results highlight the significant role of PFS and EFS as early efficacy endpoints in designing prospective trials.

As the association of improved PFS or EFS with prolonged OS in DLBCL in this study is straightforward, the use of PFS and EFS as early efficacy endpoints not only incorporates survival, but also reduces treatment-related events, disease relapse, and progression. Compared with long-term OS, dynamic assessment of PFS or EFS at 1–3 years has a lower likelihood of confounding by subsequent or salvage treatment. Innovative treatment strategies with a large magnitude of effect on PFS or EFS for high-risk patients with DLBCL may have a large effect on OS in RCTs. Importantly, we found that PFS or EFS as early as 1 year correlated with 5-year OS at the treatment arm-level, mainly because the majority of patients were at high risk of early relapse and poor post-progression survival. Consistent with this finding, other studies have demonstrated that ~70% of disease failures occurred within the first year after treatment, but rarely after 5 years [9, 12]. For patients who achieved EFS at 12 and 24 months, the risk of relapse in the next 5 years dropped to 13% and 8%, respectively [9]. If patients experienced progression or relapse within 2 years, the median OS after disease progression was only 7.2 months [11].

The strengths of this study include the quality control design, large sample size, external validation of PFS outcomes, and current standard treatment. First, the data were obtained from high-quality RCTs, phase II, and retrospective studies that enrolled large-scale cohorts (>31,000 patients) with newly diagnosed DLBCL uniformly treated with rituximab-containing immunochemotherapy. We could eliminate selection bias with great confidence due to the limited number of RCTs or treatment option heterogeneity. This comprehensive surrogacy study at trial- and treatment arm-level complements previous evidence and strengthens the clinical use of PFS and EFS as early efficacy endpoints. Second, the positive relationships between the 1–5-year PFS and 5-year OS rates were externally validated using independent data that included patients across different countries with varied eligibility criteria, immunochemotherapy regimens, radiotherapy, and follow-up times. As a variety of immunochemotherapy regimens was investigated in a heterogeneous population, we could examine for variability in treatment outcomes and hence improved the generalizability of our study. Our generation and validation of prediction models for describing the association between the 1–5-year PFS and 5-year OS rates is unique. The RCT validation in an independent cohort improved the reliability of the conclusions.

The study limitations include the lack of individual patient data and standardized definition of endpoints and follow-up assessments. First, this is a literature-based systematic review without individual patient data; therefore, patient-level surrogacy was absent. Second, precise modeling requires standardized definitions of endpoints and standardized follow-up assessments or surveillance strategies in DLBCL trials, which is infeasible to accomplish in our study. For example, while PFS was calculated from the date of randomization in RCTs, it was generally calculated from diagnosis or initial therapy in retrospective studies. In addition, EFS events typically consisted of both PFS events, as well as unplanned treatment, treatment discontinuation and toxic events as they were used to evaluate the safety, toxicity or compliance of a novel therapy. Moreover, EFS events were defined inconsistently across trials and dependent on the trial design and purpose. In clinical practice, the exact date of disease progression is difficult to determine precisely, such that the reported PFS or EFS event date was naturally dependent on the frequency and interval of two consecutive clinical visits and imaging assessments. Such an inherited heterogeneity in the interval and frequency of assessments across cannot be removed nor quantified. Third, the predicted model concluded in this study was based on findings in patients treated with anthracycline-based immunochemotherapy, and its extrapolation to other treatments would be speculative. The impact of post-progression management was beyond the scope of this study, and such information is not routinely collected in clinical trials. When more effective salvage treatment occurs and post-progression survival is significantly prolonged in the future, the predicted model should also be modified and optimized. Fourth, the correlation between EFS and OS was not externally validated in the retrospective populations, because EFS is generally not reported in retrospective studies.
In conclusion, our assessment of a large sample of high-quality data for patients with DLBCL provides high-level evidence that PFS and EFS are valid early efficacy endpoints for OS in the immunochemotherapy era.

Acknowledgements This work was supported by the CAMS Innovation Fund for Medical Sciences (CIFMS) [grant no. 2016-I2M-1-001, 2017-I2M-3-005], the National Key Projects of Research and Development of China [grant no. 2016YFC0904600], and the National Natural Science Foundation of China (grant no. 81670185). The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and the decision to submit the paper for publication.

Author contributions YXL, SNQ, and CH designed the study, analyzed the data, and revised the paper. IZ performed the literature search and quality assessment, extracted and analyzed data, and wrote the first draft of the paper. JT performed the literature search. YY, BC, SLW, and JRD analyzed data. CH supervised data analysis.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Coiffier B, Thieblemont C, Van Den Neste E, Lepeu G, Plantier I, Castaigne S, et al. Long-term outcome of patients in the LNH-98.5 trial, the first randomized study comparing rituximab-CHOP to standard CHOP chemotherapy in DLBCL patients: a study by the Groupe d’Etudes des Lymphomes de l’Adulte. Blood 2010;116:2040–5.
2. Pfreundschuh M, Kuhn T, Trumper L, Osterborg A, Trneny M, Shepherd L, et al. CHOP-like chemotherapy with or without rituximab in young patients with good-prognosis diffuse large-B-cell lymphoma: 6-year results of an open-label randomised study of the MabThera International Trial (MInT) Group. Lancet Oncol 2011;12:1013–22.
3. Habermann TM, Weller EA, Morrison VA, Gascoyne RD, Casileth PA, Cohn JB, et al. Rituximab-CHOP versus CHOP alone or with maintenance rituximab in older patients with diffuse large B-cell lymphoma. J Clin Oncol. 2006;24:3121–7.
4. Pfreundschuh M, Schubert J, Ziepert M, Schmits R, Mohren M, Lengfelder E, et al. Six versus eight cycles of bi-weekly CHOP-14 with or without rituximab in elderly patients with aggressive CD20+ B-cell lymphomas: a randomised controlled trial (RICOVER-60). Lancet Oncol 2008;9:105–16.
5. Sehn LH, Gascoyne RD. Diffuse large B-cell lymphoma: optimizing outcome in the context of clinical and biologic heterogeneity. Blood 2015;125:22–32.
6. Sargent DJ, Wieand HS, Haller DG, Gray R, Benedetti JK, Buyse M, 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 2005;23:8664–70.
7. Happold G, Zheng Y, Sehn LH, Villa D, Klasa R, Gerrie AS, et al. Evaluation of the risk of relapse in classical Hodgkin lymphoma at event-free survival time points and survival comparison with the general population in British Columbia. J Clin Oncol 2016;34:2493–500.
8. Raphael MJ, Robinson A, Booth CM, O’Donnell J, Palmer M, Eisenhauer E, et al. The value of progression-free survival as a treatment end point among patients with advanced cancer: a systematic review and qualitative assessment of the literature. JAMA Oncol. 2019;5:1779–89.
9. Maurer MJ, Ghesquières H, Jais JP, Witzig TE, Haioun C, Thompson CA, et al. Event-free survival at 24 months is a robust end point for disease-related outcome in diffuse large B-cell lymphoma treated with immunochemotherapy. J Clin Oncol 2014;32:1066–73.
10. Shi Q, Schmitz N, Ou FS, Dixon JG, Cunningham D, Pfreundschuh M, et al. 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). J Clin Oncol. 2018;36:2593–602.
11. Maurer MJ, Habermann TM, Shi Q, Schmitz N, Cunningham D, Pfreundschuh M, et al. Progression-free survival at 24 months (PFS24) and subsequent outcome for patients with diffuse large B-cell lymphoma (DLBCL) enrolled on randomized clinical trials. Ann Oncol 2018;29:1822–7.
12. Jakobsen LH, Bøgsted M, Brown PN, Arboe B, Jørgensen J, Larsen TS, et al. Minimal loss of lifetime for patients with diffuse large B-cell lymphoma in remission and event free 24 months after treatment: a Danish population-based study. J Clin Oncol 2017;35:778–84.
13. Rêcher C, Coiffier B, Haioun C, Molina TJ, Fermé C, Casasnovas O, et al. Intensified chemotherapy with ACVBP plus rituximab versus standard CHOP plus rituximab for the treatment of diffuse large B-cell lymphoma (LNH03-2B): an open-label randomised phase 3 trial. Lancet 2011;378:1858–67.
14. Schmitz N, Nickelsen M, Ziepert M, Haenel M, Borchmann P, Schmidt C, et al. Conventional chemotherapy (CHOP-14) with rituximab or high-dose chemotherapy (MegaCHOEP) with rituximab for young, high-risk patients with aggressive B-cell lymphoma: an open-label, randomised, phase 3 trial (DSHNHL 2002-1). Lancet Oncol 2012;13:1290–9.
15. Merli F, Luminari S, Rossi G, Mammì C, Marcheselli L, Tucci A, et al. Cyclophosphamide, doxorubicin, vincristine, prednisone and rituximab versus epirubicin, cyclophosphamide, vinblastine, prednisone and rituximab for the initial treatment of elderly “fit” patients with diffuse large B-cell lymphoma: results from the ANZINTER3 trial of the Intergruppo Italiano Linfomi. Leuk Lymphoma 2012;53:581–8.
16. Delarue R, Tilly H, Mounier N, Petrella T, Salles G, Thieblemont C, et al. Dose-dense rituximab-CHOP compared with standard rituximab-CHOP in elderly patients with diffuse large B-cell lymphoma (the LNH03-6B study): a randomised phase 3 trial. Lancet Oncol 2013;14:525–33.
17. Cunningham D, Hawkes EA, Jack A, Qian W, Smith P, Mouncey P, et al. Rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisolone in patients with newly diagnosed diffuse large B-cell non-Hodgkin lymphoma: a phase 3 comparison of dose intensification with 14-day versus 21-day cycles. Lancet 2013;381:1817–26.

18. Jaeger U, Trneny M, Melzer H, Praxmarer M, Nawarawong W, Ben Yehuda D, et al. Rituximab maintenance for patients with aggressive B-cell lymphoma in first remission: results of the randomized NHL13 trial. Haematologica 2015;100:955–63.

19. Seymour JF, Pfreundschuh M, Trneny M, Sehn LH, Catalano J, Csinady E, et al. R-CHOP with or without bevacizumab in patients with previously untreated diffuse large B-cell lymphoma: final MAIN study outcomes. Haematologica 2014;99:1343–9.

20. Haion C, Mounier N, Emile JF, Ranta D, Coiffrer B, Tilly H, et al. Rituximab versus observation after high-dose consolidate first-line chemotherapy with autologous stem-cell transplantation in patients with poor-risk diffuse large B-cell lymphoma. Ann Oncol 2009;20:1985–92.

21. Herbrecht R, Cernohous P, Engert A, Le Guillou S, Macdonald D, Machida C, et al. Comparison of paxitron-based regimen (CPOP-R) with doxorubicin-based therapy (CHOP-R) for treatment of diffuse large B-cell lymphoma. Ann Oncol 2013;24:2618–23.

22. Li X, Huang H, Xu B, Guo H, Lin Y, Ye S, et al. Dose-dense rituximab-CHOP versus standard rituximab-CHOP in newly diagnosed Chinese patients with diffuse large B-cell lymphoma: a randomized, multicenter, open-label phase 3 trial. Cancer Res Treat. 2019;51:919–32.

23. Chiappella A, Martelli M, Angelucci E, Brusamolino E, Evangelista A, Currella AM, et al. Rituximab-dose-dense chemotherapy with or without high-dose chemotherapy plus autologous stem-cell transplantation in high-risk diffuse large B-cell lymphoma (DLCL04): final results of a multicentre, open-label, randomised, controlled, phase 3 study. Lancet Oncol 2017;18:1076–88.

24. Cortelazzo S, Tarella C, Gianni AM, Ladetto M, Barbui AM, Rossi A, et al. Randomized trial comparing R-CHOP versus high-dose sequential chemotherapy in high-risk patients with diffuse large B-cell lymphomas. J Clin Oncol. 2016;34:4015–22.

25. Bartlett NL, Wilson WH, Jung SH, Hsi ED, Maurer MJ, Pederson LD, et al. Dose-adjusted EPOCH-R compared with R-CHOP as frontline therapy for diffuse large B-cell lymphoma: clinical outcomes of the phase III intergroup trial Alliance/CalGB 50303. J Clin Oncol. 2019;37:1790–9.

26. Poeschel V, Held G, Ziepert M, Witzig TE, Tobinai K, Rigacci L, Ikeda T, Vanazzi A, Hino M, et al. Adjuvant everolimus in high-risk diffuse large B-cell lymphoma: final results from the PILLAR-2 randomized phase III trial. Ann Oncol 2018;29:707–14.

27. Dührsen U, Müller S, Hertenstein B, Thomssen H, Kotzerke J, Mesters R, et al. Positron emission tomography-guided therapy of aggressive non-Hodgkin lymphomas (PETAL): a multicenter, randomized phase III trial. J Clin Oncol. 2018;36:2024–34.

28. Lamy T, Damaj G, Soubeyran P, Gyan E, Carton G, Bouabdallah K, et al. R-CHOP 14 with or without radiotherapy in nonbulky limited-stage diffuse large B-cell lymphoma. Blood 2018;131:174–81.

29. Vitolo U, Trneny M, Belada D, Burke JM, Carella AM, Chua N, et al. Obinutuzumab or rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone in previously untreated diffuse large B-cell lymphoma. J Clin Oncol. 2017;35:3529–37.

30. Davies A, Cummins TE, Barrans S, Maisman T, Mamot C, Novak U, et al. Gene-expression profiling of bortezomib added to standard chemoimmunotherapy for diffuse large B-cell lymphoma (REMODL-B): an open-label, randomised, phase 3 trial. Lancet Oncol 2019;20:649–62.

31. Younes A, Sehn LH, Johnson P, Zinzani PL, Hong X, Zhu J, et al. Randomized phase III trial of ibritinib and rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone in non-germinal center B-cell diffuse large B-cell lymphoma. J Clin Oncol. 2019;37:1285–95.

32. Lugtenburg P, Avivi I, Berenschot H, Ilhan O, Marolleau JP, Nagler A, et al. Efficacy and safety of subcutaneous and intravenous rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone in first-line diffuse large B-cell lymphoma: the randomized MabEase study. Haematologica 2017;102:1913–22.

33. Ketterer N, Coiffrer B, Thieblemont C, Fermé C, Brière J, Casasnovas O, et al. Phase III study of ACVBP versus ACVB plus rituximab for patients with localized low-risk diffuse large B-cell lymphoma (LHN03-1B). Ann Oncol 2012;24:1032–7.

34. Wells GA, Shea B, O’Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle–Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. 2009. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. Accessed 31 Dec 2019.

35. Casasnovas RO, Ysbaert L, Thieblemont C, Bachy E, Feugier P, Delmer A, et al. FDG-PET-driven consolidation strategy in diffuse large B-cell lymphoma: final results of a randomized phase 2 study. Blood 2017;130:1315–26.

36. Murawski N, Pfreundschuh M, Zeynalova S, Poeschel V, Hänel M, Held G, et al. Optimization of rituximab for the treatment of DLBCL (I): dose-dense rituximab in the DENSE-R-CHOP-14 trial of the DSHNHL. Ann Oncol 2014;25:1800–6.

37. Fitoussi O, Belhadj K, Mounier N, Parsons M, Tilly H, Salles G, et al. Survival impact of rituximab combined with ACVBP and upfront consolidation autotransplantation in high-risk diffuse large B-cell lymphoma for GELA. Haematologica 2011;96:1136–43.

38. Niitsu N, Kohri M, Hagiwara Y, Tanae K, Takahashi N, Bessho Y, et al. Comparison of pixantrone-based regimen CHOP-R versus CHOP-R for first-line diffuse large B-cell lymphoma: results of the cell-of-origin classiﬁcation group (CPOP-R) with doxorubicin-based therapy (CHOP-R) for treatment of diffuse large B-cell lymphoma. Ann Oncol 2009;20:1985–92.

39. Niitsu N, Kohri M, Hagiwara Y, Tanae K, Takahashi N, Bessho Y, et al. Comparison of pixantrone-based regimen CHOP-R versus CHOP-R for first-line diffuse large B-cell lymphoma: results of the cell-of-origin classiﬁcation group (CPOP-R) with doxorubicin-based therapy (CHOP-R) for treatment of diffuse large B-cell lymphoma. Ann Oncol 2009;20:1985–92.

40. Murawski N, Pfreundschuh M, Zeynalova S, Poeschel V, Hänel M, Held G, et al. Optimization of rituximab for the treatment of DLBCL (I): dose-dense rituximab in the DENSE-R-CHOP-14 trial of the DSHNHL. Ann Oncol 2014;25:1800–6.

41. Fitoussi O, Belhadj K, Mounier N, Parsons M, Tilly H, Salles G, et al. Survival impact of rituximab combined with ACVBP and upfront consolidation autotransplantation in high-risk diffuse large B-cell lymphoma for GELA. Haematologica 2011;96:1136–43.

42. Niitsu N, Kohri M, Hagiwara Y, Tanae K, Takahashi N, Bessho Y, et al. Comparison of pixantrone-based regimen CHOP-R versus CHOP-R for first-line diffuse large B-cell lymphoma: results of the cell-of-origin classiﬁcation group (CPOP-R) with doxorubicin-based therapy (CHOP-R) for treatment of diffuse large B-cell lymphoma. Ann Oncol 2009;20:1985–92.
58. Park YH, Yi HG, Lee MH, Kim CS, Lim JH. Prognostic value of TP53 in diffuse large B-cell lymphoma. J Cancer 2018;9:1231–8.

59. Song MK, Chung JS, Shin DY, Lim SN, Lee GW, Choi JC, et al. Prognostic significance of serum beta-2 microglobulin in patients with diffuse large B-cell lymphoma in the rituximab era. Oncotarget 2016;7:76934–43.

60. Tsuyama N, Sakata S, Baba S, Mishima Y, Nishimura N, Ueda K, et al. BCL2 expression in DLBCL: reappraisal of immunohistochemistry with new criteria for therapeutic biomarker evaluation. Blood 2017;130:489–500.

61. Alinari L, Gru A, Quincon C, Huang Y, Lozanski A, Lozanski G, et al. De novo CD 5+ diffuse large B-cell lymphoma: adverse outcomes with and without stem cell transplantation in a large, multicenter, rituximab treated cohort. Am J Hematol. 2016;91:395–9.

62. Prochazka KT, Melchardt T, Posch F, Schlick K, Deutsch A, Beham-Schmid C, et al. NCCN-IPI score-independent prognostic potential of pretreatment uric acid levels for clinical outcome of diffuse large B-cell lymphoma patients. Br J Cancer. 2016;115:1264–72.

63. Lee S, Hong JY, Yoon S, Yoo C, Park JH, Lee JB, et al. Prognostic significance of serum beta-2 microglobulin in patients with diffuse large B-cell lymphoma in the rituximab era. Oncotarget 2016;7:67543–43.

64. Dabaja BS, Vanderplas AM, Crosby-Thompson AL, Abel GA, Czuczman MS, Friedberg JW, et al. Radiation for diffuse large B-cell lymphoma in the rituximab era: analysis of the National Comprehensive Cancer Network lymphoma outcomes project. Cancer 2015;121:1032–9.

65. El-Galaly TC, Villa D, Alzahrani M, Hansen JW, Sehn LH, Wilson D, et al. Outcome prediction by extranodal involvement, IPI, R-IPI, and NCCN-IPI in the PET/CT and rituximab era: a Danish-Canadian study of 443 patients with diffuse-large B-cell lymphoma. Am J Hematol. 2015;90:1041–6.

66. Gong QX, Lu TX, Liu C, Wang Z, Liang JH, Xu W, et al. Prevalence and clinicopathologic features of CD30-positive de novo diffuse large B-cell lymphoma in Chinese patients: a retrospective study of 232 cases. Int J Clin Exp Pathol. 2015;8:15825–35.

67. Kumar A, Lanning MA, Zhang Z, Migliacci JC, Moskowitz CH, Zelenetz AD. Excellent outcomes and lack of prognostic impact of cell of origin for localized diffuse large B-cell lymphoma in the rituximab era. Br J Haematol. 2015;171:776–83.

68. Melchardt T, Troppan K, Weiss L, Hufnagl C, Neureiter D, Tränkenschuh W, et al. A modified scoring of the NCCN-IPI is more accurate in the elderly and is improved by albumin and jβ2-microglobulin. Br J Haematol. 2015;168:239–45.

69. Nakajima Y, Tomita N, Itabashi M, Miyashita K, Watanabe R, Miyazaki T, et al. Analysis of outcomes in patients with supra-diaphragmatic vs infra-diaphragmatic diffuse large B cell lymphoma treated with R-CHOP therapy. Leuk Res 2015;39:198–203.

70. Dabaja BS, Hess K, Shihadeh F, Podoloff DA, Medeiros LJ, Mawlawi O, et al. Positron emission tomography/computed tomography findings during therapy predict outcome in patients with diffuse large B-cell lymphoma treated with chemotherapy alone but not in those who receive consolidation radiation. Int J Radiat Oncol Biol Phys. 2014;89:384–91.

71. Mian M, Wasle I, Gameraith G, Mondello P, Melchardt T, Jäger T, et al. R-CHOP versus R-COMP: are they really equally effective? Clin Oncol 2014;26:648–52.

72. Castillo JJ, Sinclair N, Beltrán BE, Song MK, Ilic I, Leppa S, et al. Similar outcomes in Asian and Western patients with diffuse large B-cell lymphoma treated with R-CHOP therapy. Leuk Res 2013;37:86–91.

73. Hashimoto Y, Yokohama A, Saitoh A, Nakahashi H, Toyama K, Mitsu T, et al. Prognostic importance of the soluble form of IL-2 receptor (sIL-2Rα) and its relationship with surface expression of IL-2Rα (CD25) of lymphoma cells in diffuse large B-cell lymphoma treated with CHOP-like regimen with or without rituximab: a retrospective analysis of 338 cases. J Clin Exp Hematol. 2013;53:197–205.

74. Kojima M, Nishikii H, Takizawa J, Aoki S, Noguchi M, Chiba S, et al. MYC rearrangements are useful for predicting outcomes following rituximab and chemotherapy: multicenter analysis of Japanese patients with diffuse large B-cell lymphoma. Leuk Lymphoma 2013;54:2149–54.

75. Lu HJ, Huang YC, Liu CY, Hung MH, Hu MH, Wu CY, et al. Diminishing prognostic role of preexisting diabetes mellitus for...
patients with diffuse large B-cell lymphoma in the rituximab era. Ann Hematol 2013;92:1495–501.

76. Ozbalak M, Ar MC, Tuzuner N, Salihoglu A, Eskazan AE, Ongoren Aydin S, et al. Detailed analysis of diffuse large B cell lymphoma patients: a single-center, retrospective study. ISRN Hematol 2013;2013:9.

77. Shi Z, Das S, Okwan-Duodu D, Esiashvili N, Flowers C, Chen Z, et al. Patterns of failure in advanced stage diffuse large B-cell lymphoma patients after complete response to R-CHOP immunochemotherapy and the emerging role of consolidative radiation therapy. Int J Radiat Oncol Biol Phys. 2013;86:569–77.

78. Tomita N, Takasaki H, Miyashita K, Fujisawa S, Ogusa E, Matsumura S, et al. R-CHOP therapy alone in limited stage diffuse large B-cell lymphoma. Br J Haematol. 2013;161:383–8.

79. Castillo JJ, Beltran BE, Song MK, Ilic I, Leppa S, Nurmi H, et al. The Hans algorithm is not prognostic in patients with diffuse large B-cell lymphoma treated with R-CHOP. Leuk Res 2012;36:413–7.

80. Huang HH, Xiao F, Chen FY, Wang T, Li JM, Wang JM, et al. Reassessment of the prognostic value of the International Prognostic Index and the revised International Prognostic Index in patients with diffuse large B-cell lymphoma: a multicentre study. Exp Ther Med. 2012;4:475–80.

81. Li ZM, Huang JJ, Xia Y, Sun J, Huang Y, Wang Y, et al. Blood lymphocyte-to-monocyte ratio identifies high-risk patients in diffuse large B-cell lymphoma treated with R-CHOP. PLoS One 2012;7:e41658.

82. Li XY, Liu Z, Cao JN, Hong X, Wang J, Chen F, et al. Rituximab in combination with CHOP chemotherapy for the treatment of diffuse large B cell lymphoma in China: a 10-year retrospective follow-up analysis of 437 cases from Shanghai Lymphoma Research Group. Ann Hematol 2012;91:837–45.

83. Lin TL, Kuo MC, Shih LY, Dunn P, Wang PN, Wu JH, et al. The impact of age, Charlson comorbidity index, and performance status on treatment of elderly patients with diffuse large B cell lymphoma. Ann Hematol 2012;91:1383–91.

84. Tomita N, Sakai R, Fujisawa S, Fujimaki K, Taguchi J, Hashimoto C, et al. SIL index, comprising stage, soluble interleukin-2 receptor, and lactate dehydrogenase, is a useful prognostic predictor in diffuse large B-cell lymphoma. Cancer Sci 2012;103:1518–23.

85. Sehn LH, Scott DW, Chhanabhai M, Berry B, Ruskova A, Berkhahn L, et al. Impact of concordant and discordant bone marrow involvement on outcome in diffuse large B-cell lymphoma treated with R-CHOP. J Clin Oncol. 2011;29:1452–7.

86. Bari A, Marcheselli L, Sacchi S, Marcheselli R, Poizzi S, Ferri P, et al. Prognostic models for diffuse large B-cell lymphoma in the rituximab era: a never-ending story. Ann Oncol 2010;21:1486–91.

87. Ennishi D, Asai H, Maeda Y, Shinagawa K, Ikeda K, Yokoyama M, et al. Statin-independent prognosis of patients with diffuse large B-cell lymphoma receiving rituximab plus CHOP therapy. Ann Oncol 2010;21:1217–21.

88. Phan J, Mazloom A, Medeiros LJ, Zreik TG, Wogan C, Shihadeh F, et al. Benefit of consolidative radiation therapy in patients with diffuse large B-cell lymphoma treated with R-CHOP chemotherapy. J Clin Oncol. 2010;28:4170–6.

89. Scandurra M, Mian M, Greiner TC, Rancoita PM, de Campos CP, Chan WC, et al. Genomic lesions associated with a different clinical outcome in diffuse large B-cell lymphoma treated with R-CHOP-21. Br J Haematol. 2010;151:221–31.

90. Tierney JF, Stewart LA, Gherzi D, Burdett S, Sydes MR. Practical methods for incorporating summary time-to-event data into meta-analysis. Trials 2007;8:16.