European Journal of Haematology

WILEY

European Journal of Haematology

ORIGINAL ARTICLE OPEN ACCESS

A Real-World Data-Based Analysis of Prognostic Indices as Part of Trial Eligibility Criteria in Diffuse Large B-Cell Lymphoma Patients

Jelena Jelicic[1] | Karen Juul-Jensen[1] | Zoran Bukumiric[2] | Mikkel Runason Simonsen[3,4] | Michael Roost Clausen[5] | Ahmed Ludvigsen Al-Mashhadi[3,6] | Robert Schou Pedersen[7] | Christian Bjørn Poulsen[8,9] | Anne Ortved Gang[9,10] | Peter Brown[9,10] | Tarec Christoffer El-Galaly[1,3,10,11,12] | Thomas Stauffer Larsen[1,13]

[1]Department of Hematology, Odense University Hospital, Odense, Denmark | [2]Institute for Medical Statistics and Informatics, University of Belgrade, Faculty of Medicine, Belgrade, Serbia | [3]Department of Hematology, Clinical Cancer Research Center, Aalborg University Hospital, Aalborg, Denmark | [4]Department of Mathematical Sciences, Aalborg University, Aalborg, Denmark | [5]Department of Hematology, Vejle Hospital, Sygehus Lillebaelt, Vejle, Denmark | [6]Department of Hematology, Aarhus University Hospital, Aarhus, Denmark | [7]Department of Hematology, Regional Hospital Gødstrup, Denmark | [8]Department of Hematology, Zealand University Hospital, Roskilde, Denmark | [9]Department of Clinical Medicine, University of Copenhagen, Denmark | [10]Department of Hematology, Copenhagen University Hospital, Rigshospitalet, Denmark | [11]Clinical Epidemiology Division, Karolinska Institutet, Stockholm, Sweden | [12]Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden | [13]Department of Clinical Research, University of Southern Denmark, Odense, Denmark

Correspondence: Thomas Stauffer Larsen (thomas.stauffer.larsen@rsyd.dk)

Received: 21 May 2024 | **Revised:** 19 August 2024 | **Accepted:** 21 August 2024

Keywords: clinical trial | diffuse large B cell lymphoma | prognostic models | selection criteria

ABSTRACT

Objectives: Recent front-line clinical trials used the International Prognostic Index (IPI) to identify trial-eligible patients with newly diagnosed diffuse large B-cell lymphoma (DLBCL). However, many IPI-like variants with improved accuracy have been developed over the years for rituximab-treated patients.

Methods: We assessed the impact of International Prognostic Indices on patient enrolment in clinical trials, aiming to exclude low-risk IPI patients based on POLARIX/EPCORE DLBCL-2 trial criteria.

Results: We identified 2877 patients in the Danish Lymphoma Registry who would have been eligible for the POLARIX trial if patients with IPI 0–1 scores were included. IPI and NCCN-IPI assigned 33.3% and 11.9% of patients to the low-risk group, respectively. Shorter 5-year overall survival (91.4% vs. 97.5%), higher relapse rate (9.9% vs. 4.4%), and more deaths (16.1% vs. 4.4%) occurred in the low-risk IPI group compared with low-risk NCCN-IPI group. Analyzed models failed to identify true high-risk patients with poor prognosis. Similar results were found in the confirmatory cohort developed based on EPCORE DLBCL-2 trial eligibility criteria.

Conclusion: True low-risk patients are more optimal identified by NCCN-IPI and should be excluded from front-line clinical trials due to their excellent prognosis. However, additional high-risk factors besides clinical prognostic models need to be considered when selecting trial-eligible patients.

1 | Introduction

The most commonly used prognostic index for risk stratification in patients with diffuse large B-cell lymphoma (DLBCL) is the International Prognostic Index (IPI) [1]. This model was developed in 1993 for patients with aggressive lymphoma treated without rituximab [2]. IPI stratifies patients into four risk groups using five easily accessible markers (age, Ann

Arbor stage, number of extranodal sites, Eastern Oncology Cooperative Group performance status [ECOG PS], and lactate dehydrogenase [LDH]) [2]. However, with the introduction of rituximab, the prognostic value of the IPI has been challenged [1, 3–5]. Therefore, several IPI-like variants have been developed over the years, with revised IPI (R-IPI) and the National Comprehensive Cancer Network IPI (NCCN-IPI) being validated in different populations, including patients treated in clinical trials [1].

Over the past 20 years, several first-line randomized controlled trials (e.g., MAIN, PIX203, ALLIANCE/CALGB50303, and ReMoDL-B) have incorporated experimental agents in the treatment of newly diagnosed DLBCL patients with R-CHOP (rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone) as control arm, but have demonstrated negative results [6–9]. These first-line studies did not select patients based on risk models and may presumably have motivated the criteria for patient selection in the most recent trials. However, these negative results could be somewhat attributed to the traditional eligibility criteria, which have become increasingly restrictive [10]. Additionally, patients with more aggressive disease and shorter intervals between diagnosis and treatment are more likely to be excluded due to the need for immediate treatment [11]. Therefore, Harkins et al. proposed modernized eligibility criteria based on expert opinion to facilitate the enrollment of patients in the first-line randomized controlled trials. Regarding prognostic models, the authors recommend including the IPI score or elements of the IPI score in trial eligibility criteria [12].

Some recent clinical trials used prognostic scores as part of eligibility criteria in patients with newly diagnosed DLBCL with R-CHOP as the control arm. In the phase 3 open-label POLARIX study, comparing polatuzumab vedotin plus R-CHP (rituximab, cyclophosphamide, doxorubicin, and prednisone) versus R-CHOP, participants were required to have IPI score ≥2. This trial achieved its primary efficacy endpoint (progression-free survival [PFS]), making it the only positive trial so far [13]. The ongoing phase 3 trial with epcoritamab plus R-CHOP versus R-CHOP also used IPI ≥2 as inclusion criteria (EPCORE DLBCL-2 and NCT05578976). IPI and age-adjusted IPI (aaIPI) are part of the inclusion criteria for the phase 3 study investigating a combination of Tafasitamab, Lenalidomide, and R-CHOP versus R-CHOP in newly diagnosed high-intermediate and high-risk DLBCL patients (frontMIND and NCT04824092). Other recent open-label randomized phase 3 trials (e.g., GOYA, PHOENIX, and LNH03-2B), all negative trials, have also used International Prognostic Indices (IPI, R-IPI, and aaIPI) as part of eligibility criteria with additional trial requirements (e.g., age, bulky disease, and cell of origin) [14–16]. Although many first-line clinical trials mainly aim to identify high-risk patients, different eligibility criteria have been used, impacting the ability to generalize results of modern clinical trials to real-world populations [10]. Trial-eligible patients categorized as low-risk by various models are excluded based on their projected good prognosis and potential concerns regarding added toxicity. However, some of the excluded patients with unfavorable outcomes on R-CHOP have high-risk characteristics that are not captured by traditional prognostic models [17]. Although more restrictive eligibility criteria lead to the recruitment of a more homogenous population, increasing effect size, such an approach limits patients' accessibility to clinical trials and the applicability of results to broader clinical populations [10].

We conducted a study investigating the impact of International Prognostic Indices (IPI, R-IPI, and NCCN-IPI) as a selection criterion for entering clinical trials. We used the POLARIX trial inclusion criteria as a template for the study and aimed to determine the number of patients from a Danish population-based registry who would be excluded if models other than IPI were used to select patients for these clinical trials. The results were tested by employing a similar strategy using eligibility criteria for the EPCORE DLBCL-2 clinical trial.

2 | Methods

The initial search included all adult patients with newly diagnosed DLBCL between 2000 and 2021, identified through the Danish lymphoma registry (LYFO) [18]. The period was selected because the database has been highly complete since the early 2000s, allowing for long-term follow-up and capturing events occurring late in the follow-up period.

Only patients who could tolerate immuno-chemotherapy were included to ensure the identification of patients eligible for curative intended treatments. After identifying patients in LYFO, we created a LYFO POLARIX cohort by including patients who met approximated POLARIX inclusion criteria. The inclusion criteria were based on the phase 3 first-line trial (POLARIX trial), which enrolled newly diagnosed DLBCL patients with IPI 2–5.

Similarly, another independent cohort (LYFO EPCORE DLBCL-2) was formed using trial criteria from a phase 3 first-line DLBCL trial (EPCORE DLBCL-2), which also included patients with IPI 2–5. This was done to investigate whether results from the POLARIX cohort could be generalized to other trials using the same inclusion criteria based on the prognostic model.

Two independent cohorts (LYFO POLARIX and LYFO EPCORE DLBCL-2) were established by retrieving the trial inclusion criteria and selecting patients according to variables available in LYFO based on criteria corresponding to the respective trials (Table S1). Trial criteria included clinical and laboratory data, which were similar between the two trials and included the following: age, absolute neutrophil count (ANC), hemoglobin, platelet count, serum alanine transaminase, total bilirubin level, and creatinine clearance. Moreover, patients with ECOG PS >2, central nervous system involvement, and a history of other prior malignancies (excluding nonmelanoma skin cancer) were excluded from the analysis.

Although data on cardiovascular disease were not provided in LYFO, only patients treated with standard R-CHOP or intensified regimes with the addition of etoposide (R-CHOEP/R-EPOCH) were included in the analysis. Patients treated with dose-reduced regimens such as R-miniCHOP and regimens without doxorubicin or similar were excluded, and this criterion served as an approximation for either heart disease or lack of fitness for standard treatment and inclusion in clinical trials. Additionally, as ANC is not registered in LYFO, this variable was approximated by subtracting lymphocyte count

from total leucocyte count. Moreover, as LYFO only provides data on creatinine, renal function was estimated using the CKD-EPI formula for the estimated glomerular filtration rate (eGFR) [19]. Treatment response was registered as complete remission (CR), partial remission (PR), stable disease (SD), progressive disease (PD), and death (if patients died before response evaluation was performed), and not evaluated (NE) if data were missing [20].

The article refers to patients in the current study as the LYFO POLARIX cohort and LYFO EPCORE DLBCL-2 cohort to highlight the difference between the original POLARIX and EPCORE DLBCL-2 trial cohorts.

2.1 | Statistical Analysis

Overall survival (OS) was estimated from diagnosis until death from any cause or censoring on March 31, 2022, while PFS represented the time from diagnosis until any of the following events: relapse, disease progression, death, or censoring at the last follow-up (March 31, 2022). OS was estimated using the Kaplan–Meier estimator, and the log-rank test was used to test for differences in survival between risk groups.

We used the concordance index to measure discrimination, with 1 representing perfect discrimination and 0.5 no discrimination [21, 22]. Akaike's information criterion (AIC) was used as a measure of fitness with a difference of ≥10, indicating a substantial improvement in the fit of the model [23]. Interrater-weighted κ statistic with 95% confidence intervals (CI) was used to compare the agreement between the IPI and NCCN-IPI [24].

All tests were two-sided, and p values below 0.05 were considered statistically significant. Statistical analyses were performed in IBM SPSS statistics (version 28.0.0.0) and R version 3.4.1 using the following packages for survival and performance calculations: CPE, ggplot2, ggsurvfit, dynpred, maxstat, survC1, and survival.

3 | Results

3.1 | LYFO POLARIX Cohort

The initial search in LYFO identified 6074 newly diagnosed DLBCL in the inclusion period. Of those, 5323 patients fulfilled the age criterion (18–80 years) for the LYFO POLARIX part of the study, and 2877 patients (median age 65 years) were included in the final analysis after the trials' inclusion/exclusion criteria were applied (Figure 1). The basic clinical characteristics of analyzed patients are presented in Table 1.

3.1.1 | Prognostic Models and Model Agreement

Table 2 provides the distribution of patients according to IPI, NCCN-IPI, and R-IPI. IPI classified 959 (33.3%) patients as low-risk (score 0–1). In contrast, NCCN-IPI allocated 343 (11.9%) patients to the low-risk group (score 0–1). R-IPI identified 281 (9.8%) as low-risk, equivalent to an IPI score of 0 (Table 2). According to both IPI and NCCN-IPI, 338 patients were identified as low-risk, while the remaining IPI low-risk patients (621 patients) were allocated to the NCCN-IPI low- and low-intermediate groups. Figure 2 provides a graphical presentation of patient distribution and mortality reclassification table. When NCCN-IPI was used as the reference model, it showed substantial agreement (weighted $\kappa = 0.61$) with IPI.

3.1.2 | Treatment Response and Outcome

A total of 2463 patients (85.6%, 95% CI, 84.3–86.8) achieved CR, while 382 (13.3%, 95% CI, 12.1–14.6) had an unfavorable response (PR/SD/PD/death) and for 32 patients (1.1%) response data were not available. Detailed information on treatment outcomes and early death across risk groups is provided in Table 2. During follow-up (5.8 years), 484 patients (16.8%, 95% CI, 15.5–18.2) relapsed, and 831 died (28.9%, 95% CI, 27.4–30.7).

When patients with IPI scores 0–1 were excluded, CR was achieved in 1572/1918 patients (82%, 95% CI, 80.2–83.6), and 321 (16.7%, 95% CI, 15.1–18.5) patients had unfavorable response. In patients with IPI scores 0–1, CR was achieved in 891 (92.9%, 95% CI, 90.0–93.4), while 95 (9.9%) relapsed. Twelve patients died within 6 months following diagnosis. However, these patients were older (median 76 years, range 57–80) than the calculated median age of 57 for the low-risk IPI group. In total, 154 patients (16.1%, 95% CI, 13.7–18.3) died in the IPI low-risk group (Table 2).

In patients with the NCCN-IPI score 0–1, CR was achieved in 329 (95.9%, 95% CI, 93.3–97.5), while only 13 (3.8%, 95% CI, 2.2–6.4) did not achieve CR, with 15 patients relapsing and dying during follow-up (Table 2). When excluding patients with low-risk NCCN-IPI, 2134 of 2534 patients achieved CR (84.2%, 95% CI, 82.7–85.6), and 369 obtained unfavorable response (14.6%, 95% CI, 13.2–16.0).

Regarding 621 of 951 low-risk IPI patients reclassified into a higher category according to NCCN-IPI, CR was achieved in 567 (91.3%) patients. During follow-up, a higher proportion of patients in this group relapsed (13.0% vs. 9.9%) and died (22.4% vs. 16.1%) compared with the low-risk IPI group.

3.1.3 | Survival Analysis

The median follow-up for surviving patients in the LYFO POLARIX cohort was 6.7 years (interquartile range [IQR] 3.9–10.3). When patients with low-risk IPI were excluded, the median follow-up was slightly shorter (5.1 years). In contrast, median follow-up remained almost the same with the exclusion of low-risk NCCN-IPI patients (5.6 years). The median OS was not reached for the entire LYFO POLARIX cohort population. When the IPI low-risk group was excluded, the median OS was 13.2 years (95% CI, 11.9–14.4), and with the exclusion of the NCCN-IPI low-risk group, the median OS was 14.1 years (95% CI, 13.2–15.0).

Figure 3 shows OS curves for all three models investigated in LYFO POLARIX. The 3- and 5-year PFS/OS rates for all models are shown in Table S2. Five-year OS estimates in the respective low-risk groups in the LYFO POLARIX cohort were 91.4%, 97.5%, and 97.9% for IPI, NCCN-IPI, and R-IPI

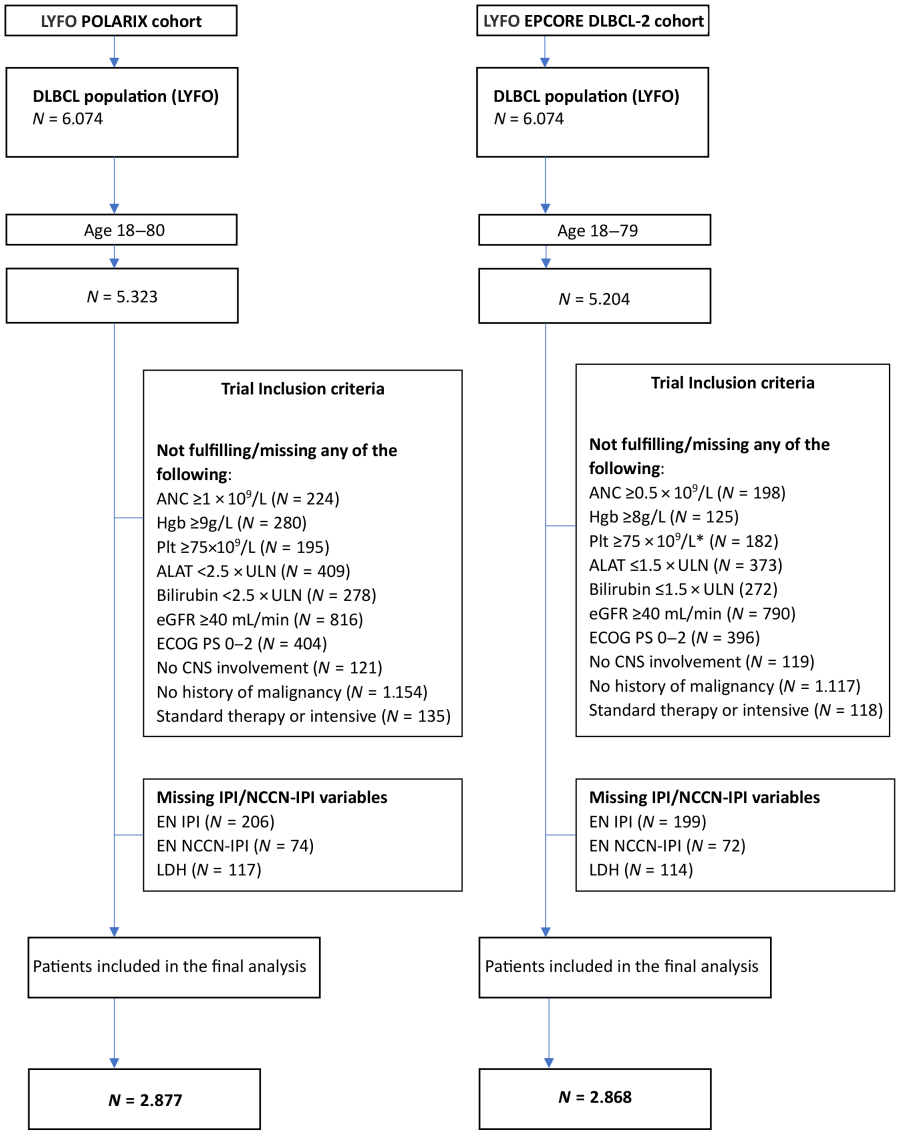

FIGURE 1 | Consort diagram of the selection process for identifying patients eligible for both LYFO POLARIX and EPCORE DLBCL-2 cohorts. *Or $\geq 25 \times 10e^9$/L in the presence of bone marrow involvement. ALAT, alanine transaminase; ANC, absolute neutrophil count; CNS, central nervous system; ECOG PS, Eastern Oncology Cooperative Group performance status; eGFR, estimated glomerular filtration rate; EN, extranodal sites; Hgb, hemoglobin; IPI, International Prognostic Index; LDH, lactate dehydrogenase; *N*, number; NCCN-IPI, National Comprehensive Cancer Network IPI; Plt, platelet count; ULN, upper limit of normal.

and 47.3% for high-risk NCCN-IPI, 56.2% for IPI, and 67.5% for R-IPI (Table S2).

Figure 4 displays the survival curves of low-risk IPI patients reclassified by NCCN-IPI and R-IPI. Patients equally classified as low-risk by IPI and NCCN-IPI (very low-risk—VLR) had a 5-year OS of 97.5%. However, patients classified as low-risk by IPI but higher-risk by NCCN-IPI (high low-risk—HLR) had a poorer 5-year OS of 88.1% than patients identified as low-risk by IPI. The respective groups for R-IPI also showed similar 5-year OS (97.9% vs. 88.6%) for VLR and HLR.

The median PFS for the LYFO POLARIX cohort was not reached, while the median PFS estimate when the IPI low-risk group was excluded was 12.8 years (95% CI, 11.7–14.2). Longer median PFS was observed when excluding the NCCN-IPI low-risk group (13.9 years, 95% CI, 12.8–15.0).

3.1.4 | Model fit and Discrimination

Regarding the LYFO POLARIX cohort, the lowest AIC was registered for NCCN-IPI (11983) (Table S2). The highest c-index was estimated for NCCN-IPI (c-index = 0.653) and was statistically significant compared to IPI (c-index = 0.620, $p < 0.05$) (Table S2).

3.2 | LYFO EPCORE DLBCL-2 Cohort

Of the 6074 newly diagnosed DLBCL cases identified in the LYFO during the inclusion period, 5204 patients were aged 18–79. Among them, 2868 (median age 67 years) fulfilled EPCORE DLBCL-2 trial criteria, excluding IPI criteria (Figure 1). Despite minor discrepancies in laboratory inclusion criteria compared with the LYFO POLARIX cohort, both

TABLE 1 | Clinical characteristics of patients with diffuse large B-cell lymphoma in the LYFO POLARIX cohort.

Clinical characteristics (LYFO POLARIX cohort)	All patients	Patients with low IPI excluded	NCCN-IPI low-risk excluded	R-IPI low-risk excluded
No. of patients	2877	1918	2534	2596
% of primary cohort ($N = 5323$ patients)	54.0	36.0	47.0	48.8
Median age	65	67	67	66
	N (%)	*N* (%)	*N* (%)	*N* (%)
Age				
≤40	225 (7.8)	92 (4.8)	96 (3.8)	156 (6.0)
41–60	834 (29.0)	385 (20.1)	620 (24.5)	622 (24.0)
61–75	1453 (50.5)	1150 (60.0)	1453 (57.3)	1453 (56.0)
>75	365 (12.7)	291 (15.2)	356 (14.4)	365 (14.1)
Gender				
Male	1693 (58.8)	1118 (58.3)	1476 (58.2)	1509 (58.1)
Female	1184 (41.2)	800 (41.7)	1058 (41.8)	1087 (41.9)
Ann Arbor stage				
I	558 (19.4)	105 (5.5)	367 (14.5)	379 (14.6)
II	496 (17.2)	124 (6.5)	361 (14.2)	394 (15.2)
III	540 (18.8)	468 (24.4)	532 (21.0)	540 (20.8)
IV	1283 (44.6)	1221 (63.7)	1274 (50.3)	1283 (49.4)
ECOG PS				
0	1826 (63.5)	1019 (53.1)	1522 (60.1)	1571 (60.5)
1	843 (29.3)	693 (36.1)	804 (31.7)	817 (31.5)
2	208 (7.2)	206 (10.7)	208 (8.2)	208 (8.0)
LDH				
<ULN	1441 (50.1)	636 (33.2)	1137 (44.9)	1160 (44.7)
1-3xULN	1234 (42.9)	1282 (56.8)	1195 (47.2)	1234 (47.5)
>3xULN	202 (7.0)	193 (10.1)	202 (8.0)	202 (7.8)
EN (IPI)				
≤1	2089 (72.6)	1141 (59.5)	1762 (69.5)	1808 (69.6)
>1	788 (27.4)	777 (40.5)	772 (30.5)	788 (30.4)
EN (NCCN-IPI)				
≤1	2052 (71.3)	1152 (60.1)	1711 (67.5)	1778 (68.5)
>1	825 (28.7)	766 (39.9)	823 (32.5)	818 (31.5)

Abbreviations: ECOG-PS, Eastern Oncology Cooperative Group performance status; EN, extranodal sites; IPI, International Prognostic Index; LDH, lactate dehydrogenase; LYFO, Danish Lymphoma Registry; NCCN-IPI, National Comprehensive Cancer Network IPI; OS, overall survival; PFS, progression-free survival; R-IPI, revised IPI.

cohorts included similar patient numbers with significant overlap (Table S3).

When EPCORE DLBCL-2 eligibility criteria were used to form the LYFO EPCORE DLBCL-2 cohort, almost identical distributions according to IPI, NCCN-IPI, and R-IPI were observed to that of the LYFO POLARIX cohort (Table S4). Regarding the low-risk group, IPI and NCCN-IPI identically identified 338 low-risk patients, while 602 low-risk IPI patients were allocated to the NCCN-IPI low- and low-intermediate group. A graphical presentation of patient distribution is provided in Figure S1.

TABLE 2 | Distribution of patients according to treatment response, relapse, and death outcome within each risk group of four risk models in the LYFO POLARIX cohort.

Models	No. of patients (%)	CR *N* (%)	PR *N* (%)	SD/PD *N* (%)	Death[a] *N* (%)	NE *N* (%)	Relapse *N* (%)	All deaths *N* (%)	Death at 30 days *N* (%)	Death at 90 days *N* (%)	Death at 180 days *N* (%)
IPI											
L	959 (33.3)	891 (93.0)	34 (3.5)	17 (1.7)	10 (1.0)	7 (0.7)	95 (9.9)	154 (16.1)	0 (0.0)	5 (0.5)	12 (1.3)
LI	785 (27.3)	683 (87.0)	44 (5.6)	31 (4.0)	18 (2.3)	9 (1.1)	130 (16.6)	217 (27.6)	1 (0.1)	10 (1.3)	25 (3.2)
HI	754 (26.2)	622 (82.5)	56 (7.4)	38 (5.0)	29 (3.8)	9 (1.2)	163 (21.6)	270 (35.8)	6 (0.8)	17 (2.3)	36 (4.8)
H	379 (13.2)	267 (70.4)	28 (7.4)	43 (11.4)	34 (9.0)	7 (1.8)	96 (25.3)	190 (50.1)	5 (1.3)	23 (6.1)	40 (10.6)
NCCN-IPI											
L	343 (11.9)	329 (95.9)	11 (3.2)	1 (0.3)	1 (0.3)	1 (0.3)	15 (4.4)	15 (4.4)	0 (0.0)	0 (0.0)	1 (0.3)
LI	1288 (44.8)	1163 (90.4)	55 (4.3)	38 (2.9)	20 (1.6)	12 (0.9)	200 (15.5)	302 (23.4)	1 (0.1)	11 (0.9)	26 (2.0)
HI	1056 (36.7)	855 (81.0)	78 (7.4)	62 (5.8)	48 (4.5)	13 (1.2)	225 (21.3)	406 (38.4)	7 (0.7)	28 (2.7)	61 (5.8)
H	190 (6.6)	116 (61.1)	18 (9.5)	28 (14.7)	22 (11.6)	6 (3.2)	44 (23.2)	108 (56.8)	4 (2.1)	16 (8.4)	25 (13.2)
NCCN-IPI[b]											
VLR	338 (35.2)	324 (95.9)	11 (3.3)	1 (0.3)	1 (0.3)	1 (0.3)	14 (4.1)	15 (4.4)	0 (0.0)	0 (0.0)	1 (0.3)
HLR	621 (64.8)	567 (91.3)	23 (3.7)	16 (2.6)	9 (1.4)	6 (1.0)	81 (13.0)	139 (22.4)	1 (0.2)	5 (0.8)	11 (1.8)
R-IPI											
L	281 (9.8)	277 (98.6)	2 (0.7)	1 (0.4)	1 (0.4)	0 (0.0)	14 (5.0)	13 (4.6)	0 (0.0)	0 (0.0)	0 (0.0)
I	1463 (50.8)	1297 (88.7)	76 (5.2)	47 (3.2)	15 (1.0)	28 (1.9)	211 (14.4)	358 (24.5)	1 (0.1)	15 (1.0)	37 (2.5)
H	1133 (39.4)	889 (78.5)	84 (7.4)	81 (7.2)	16 (1.4)	63 (5.6)	259 (22.9)	460 (40.6)	11 (1.0)	40 (3.5)	76 (6.7)
R-IPI[b]											
VLR	281 (29.3)	277 (98.6)	2 (0.7)	1 (0.4)	0 (0.0)	1 (0.4)	14 (5.0)	13 (4.6)	0 (0.0)	0 (0.0)	0 (0.0)
HLR	678 (70.7)	614 (90.6)	32 (4.7)	16 (2.4)	10 (1.5)	6 (0.9)	81 (11.9)	414 (20.8)	1 (0.1)	5 (0.7)	12 (1.8)

Abbreviations: CR, complete remission; H, high; HI, high-intermediate; HLR, high low-risk group; IPI, International Prognostic Index; L, low; LI, low-intermediate; LYFO, Danish Lymphoma Registry; NCCN-IPI, National Comprehensive Cancer Network IPI; NE, not evaluated; OS, overall survival; PD, progressive disease; PFS, progression-free survival; PR, partial remission; R-IPI, revised IPI; SD, stable disease; VLR, very low-risk.
[a]Death coded as treatment response.
[b]Conditional on a low-risk IPI.

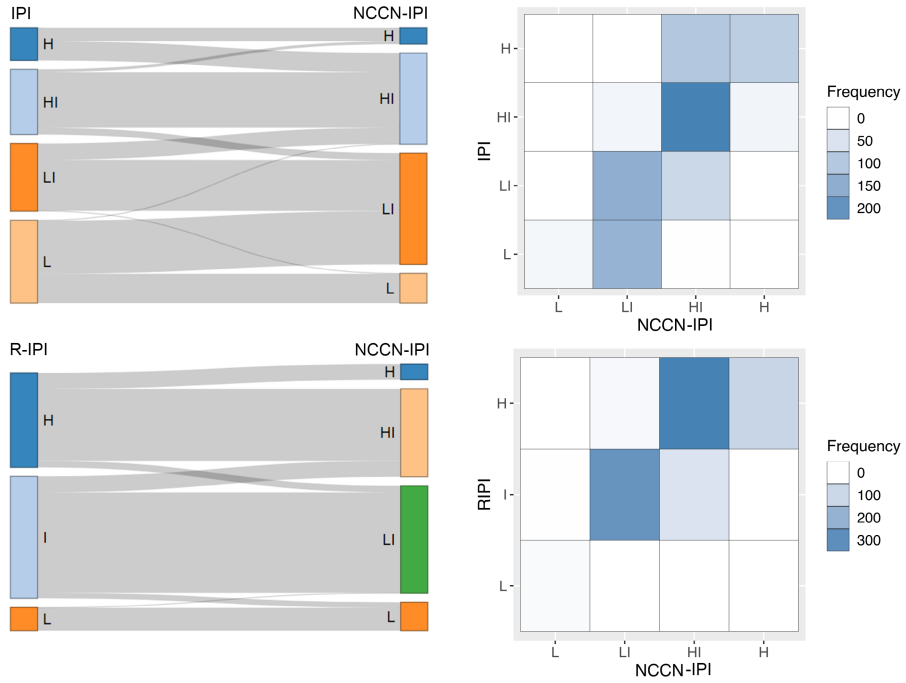

FIGURE 2 | Sankey diagram of risk groups (left colon) and reclassification mortality frequencies (right colon) with NCCN-IPI as referent model in the LYFO POLARIX cohort. H, high risk; HI, high-intermediate risk; IPI, International Prognostic Index; L, low-risk; LI, low-intermediate risk; NCCN-IPI, National Comprehensive Cancer Network IPI; R-IPI, revised IPI.

Treatment response and outcome results in the LYFO EPCORE DLBCL-2 cohort are presented in Table S4. The mortality reclassification table with NCCN-IPI as a referent model is graphically presented in Figure S1.

Regarding survival analysis, the median follow-up duration of surviving patients was identical to the POLARIX study for the entire cohort (6.7 years). This similarity also holds regarding survival estimates for the whole study population and when low-risk groups of IPI and NCCN-IPI are excluded. Figure S2 shows OS curves for all three models investigated in the EPCORE DLBCL-2 cohort. The 5-year OS estimates in the low-risk group were 91.5%, 97.8%, and 98.3% for IPI, NCCN-IPI, and R-IPI, and 48.4% for high-risk NCCN-IPI, 56.0% for IPI and 68.0% for R-IPI (Table S5). The same table also presents corresponding 3- and 5-year PFS. Moreover, NCCN-IPI (c-index = 0.655) demonstrated better discrimination between risk groups than IPI (c-index = 0.623) (Table S5).

Figure S3 displays the survival curves of low-risk IPI patients reclassified by NCCN-IPI and R-IPI. Patients equally classified as low-risk by IPI and NCCN-IPI (very low-risk—VLR) had a 5-year OS of 97.8%. However, patients classified as low-risk by IPI but higher-risk by NCCN-IPI (high low-risk—HLR) had a poorer 5-year OS of 88.0% than patients identified as low-risk by IPI. The respective groups for R-IPI also showed similar 5-year OS (98.3% vs. 88.5%) for VLR and HLR.

Regarding PFS, the median was not reached, while the median PFS estimate when the IPI low-risk group was excluded was 12.9 years (95% CI, 11.5–14.3). Longer median PFS was observed when excluding the NCCN-IPI low-risk group (14.1 years, 95% CI, 12.9–15.2).

4 | Discussion

We have used a population-based lymphoma registry to compare the impact of International Prognostic Indices as part of inclusion criteria in clinical trials. We have selected patients according to approximated POLARIX trial eligibility criteria and confirmed results in a cohort developed using approximated EPCORE DLBCL-2 trial eligibility criteria. These two front-line trials were chosen due to similar inclusion and exclusion criteria, namely low IPI risk (0–1 points). In our study, NCCN-IPI and R-IPI could successfully identify a subgroup of low-risk IPI patients at increased risk of treatment failure. Additionally, if NCCN-IPI or R-IPI was used to select patients instead of IPI, a significantly larger number of patients would be eligible for clinical trials, and patients with excellent prognosis would remain in the low-risk group and thus would have been excluded.

It is reported that up to 50% of DLBCL patients do not fulfill inclusion criteria in most recent clinical trials, with up to a quarter excluded based on organ function alone [25, 26]. Regarding our primary cohort ($n = 5323$), when all trial inclusion and exclusion criteria required by the POLARIX trial available in LYFO were used apart from IPI, 54% of all patients ($n = 2877$) in the relevant age group fulfilled the eligibility criteria. These numbers could be slightly higher as some patients were excluded due to missing data, particularly regarding renal function and prior malignancy. However, only 36% of the primary cohort were included in the LYFO POLARIX cohort if an IPI score ≥2 was used as an inclusion criterion per the original trial design. If NCCN-IPI and R-IPI were used as selection criteria, a significantly higher proportion of patients were trial-eligible. With NCCN-IPI as inclusion criteria instead of IPI, 616 more patients could enter the LYFO POLARIX cohort. Consequently, with the exclusion of

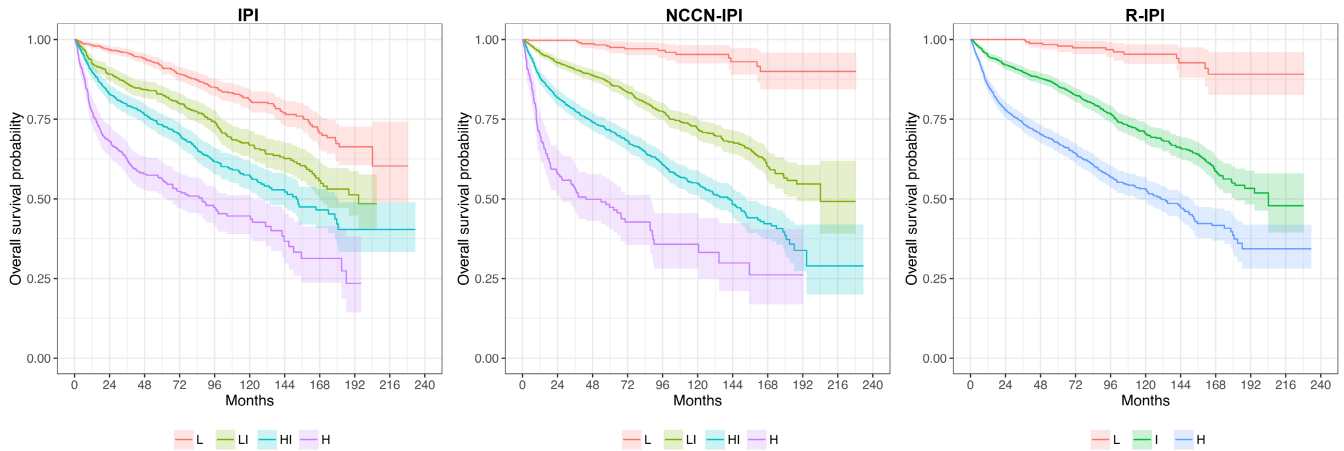

FIGURE 3 | Overall survival regarding four International Prognostic Indices in the LYFO POLARIX cohort (Kaplan–Meier curves). The shaded color areas around curves represent confidence intervals. H, high risk; HI, high-intermediate risk; IPI, International Prognostic Index; L, low-risk; LI, low-intermediate risk; NCCN-IPI, National Comprehensive Cancer Network IPI; R-IPI, revised IPI.

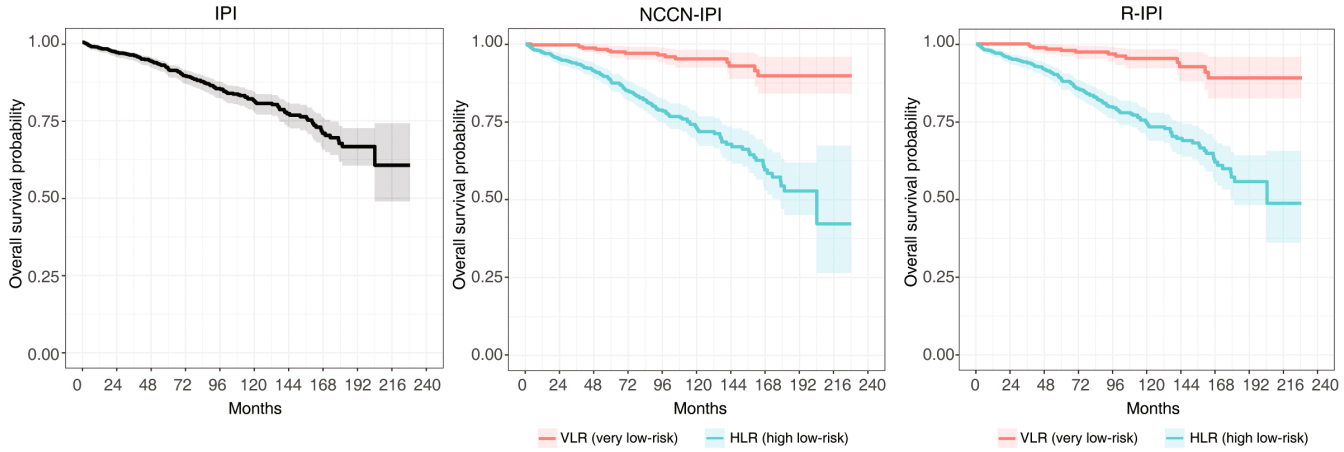

FIGURE 4 | Overall survival of low-risk IPI patients and reclassified by NCCN-IPI and R-IPI into very low-risk (VLR) and high low-risk (HLR) groups in the LYFO POLARIX cohort. IPI, International Prognostic Index; NCCN-IPI, National Comprehensive Cancer Network IPI; R-IPI, revised IPI.

low-risk NCCN-IPI, 47% of the initial population could be candidates for the respective trial. These results demonstrate that a significant fraction of the population is trial-ineligible. The findings were confirmed in the independent LYFO EPCORE DLBCL-2 cohort despite a slight difference in trial eligibility criteria between the POLARIX and EPCORE DLBCL-2 trials.

IPI identified significantly more patients as low-risk (33.1%) than R-IPI (9.8%) and NCCN-IPI (11.9%) in the LYFO POLARIX cohort. These results mimicked original data from the publications behind IPI and R-IPI [2, 3]. However, we identified a lower percentage of patients in the low-risk NCCN-IPI group compared with the study of Zhou et al., which is probably because NCCN-IPI was developed in real-life populations, while IPI was developed based on clinical trial cohorts [4]. Regarding treatment response, most patients in the current study achieved CR (85.6%). Compared with the original POLARIX trial, in which 78% in the Pola-R-CHP group and 74% in the R-CHOP group achieved CR, we observed more favorable responses also when patients with low-risk IPI (82%) and NCCN-IPI (84.2%) were excluded [13]. In our study, the improved treatment response may result from more intensive treatments in some cases and a slight risk of misclassification of certain DLBCL patients in the LYFO registry, as further explained in the limitations section. However, response rates in our study are in accordance with previously reported data on treatment responses in DLBCL patients [27].

Regarding survival, all models could discriminate between risk groups in rituximab-treated patients. Although Ruppert et al. showed that NCCN-IPI exhibited the best prognostic quality measures, the authors concluded that all three scoring systems failed to identify a high-risk group with poor long-term OS, as 5-years OS ranged from 49% to 53.9%, and 60.9% for NCCN-IPI, IPI, and R-IPI, respectively [1]. As we included trial candidates from the real-world population, we, as anticipated, found similar 5-year OS in the high-risk patients according to NCCN-IPI (47.3%), IPI (56.2%), and R-IPI (67.5%). This contrasts with the original publication behind NCCN-IPI, where the 5-year OS for high-risk NCCN-IPI and IPI were 33% and 54%, respectively. However, NCCN-IPI was developed from a real-world population, while Ruppert et al. analyzed patients treated in clinical

trials, where better outcomes in high-risk populations compared with a real-life population are somewhat expected [1, 4].

Furthermore, NCCN-IPI, in our study, could identify patients with an excellent prognosis. The 5-year OS rate in the low-risk group was 97.5%, similar to R-IPI but differed from the IPI (91.4%). We also observed more deaths in the low-risk IPI group (16.1%) in the follow-up compared with 4.4% and 4.6% in the low-risk NCCN-IPI and R-IPI. Moreover, higher relapse rates were estimated for low-risk IPI (9.9%) compared with low-risk NCCN-IPI (4.4%) and R-IPI (5.0%) groups. Most importantly, both NCCN-IPI and R-IPI could identify a subgroup of patients within the low-risk IPI population that has a greater likelihood of relapse, treatment failure, early death, and overall mortality. This result indicates that IPI did not accurately identify a low-risk population of trial-eligible patients who may experience treatment failure with standard therapy. Loss of information due to the dichotomization of variables across IPI scores may cause previously described outcome heterogeneity within IPI. Consequently, including low-risk patients with excellent prognoses can lead to potentially underpowered trials [28]. This may be prevented by excluding patients with low-risk NCCN-IPI or R-IPI instead of IPI from front-line clinical trials while balancing experimental first-line treatments' potential benefits and side effects. As clinical trials are expensive for sponsors, identifying a population that most likely benefits from experimental drugs is crucial, which could be why many front-line DLBCL clinical trials have more restrictive eligibility criteria [10]. However, more strict clinical eligibility criteria can prevent the approval of beneficial experimental treatment in specific groups of patients not included in original clinical trials [10].

Selecting eligible patients for clinical trials using prognostic models should be done carefully, considering additional emerging prognostic markers that can help to increase statistical power when designing clinical trials [17]. Many prognostically significant laboratory markers (e.g., hemoglobin, albumin, beta 2-microglobulin, and white blood cell counts) have improved discriminatory ability when added to current models [29–31]. Integrating more precise metrics measuring tumor burden (e.g., total metabolic tumor volume and the maximum distance among lesions), interim response assessments, and quantifying the tumor using circulating tumor DNA to identify patients at risk of treatment failure has shown promising results [17]. Moreover, identifying multiple genetic subtypes within and beyond the cell of origin classification could help select trial-eligible patients while testing experimental agents, allowing a more personalized approach. Last, innovative trial designs are necessary to identify a broader population of DLBCL patients who could benefit from agents under evaluation [26].

Although this is one of the most extensive studies based on real-world population data investigating the impact of International Prognostic Indices as a selection criterion for entering clinical trials, several limitations should be addressed. The retrospective nature of the study using register-based data comes with a selection bias, as around 17% of patients were excluded due to a lack of data on required variables when adjusted for trial inclusion criteria. Moreover, there is a small risk of incorrect disease classification, such as primary mediastinal B-cell lymphoma and primary effusion lymphoma, which are typically excluded from clinical trials. However, due to the rarity of these subtypes, the impact is likely to be insignificant. Nevertheless, by using inclusion criteria according to selected clinical trials, despite the necessity for additional adjustments regarding ANC, renal function, heart disease, and fitness, we were able to select a population of patients that is likely to be included in clinical trials. The study's main strength is the large number of patients from the real world included and the possibility of direct comparison of prognostic models as inclusion criteria for entering clinical trials.

5 | Conclusion

In this retrospective registry-based study, we evaluated the impact of International Prognostic Indices on selecting patients for clinical trial inclusion. Even if IPI was not part of the trial enrolment criteria, only about 55% of the registry population could be trial-eligible based on eligibility criteria from two large clinical trials. However, if patients with a low-risk IPI were excluded, over one-third of the primary eligible patients would be ineligible per the original protocol, further limiting patients eligible for a clinical trial. Substituting low-risk IPI with low-risk NCCN-IPI or R-IPI could significantly reduce the number of excluded patients, approximately 12% of our trial-eligible population. Moreover, NCCN-IPI and R-IPI identified a subgroup of patients with low-risk IPI and a less favorable prognosis. Due to a highly favorable prognosis, patients with low-risk NCCN-IPI or R-IPI should be excluded from upfront clinical trials, as this approach would spare a population with an excellent prognosis from the potential harm of experimental treatment. However, current prognostic models fail to identify high-risk patients with poor long-term survival. Therefore, identifying additional prognostic parameters besides prognostic models for selecting trial-eligible patients at risk of treatment failure is warranted.

Author Contributions

J.J., Z.B., K.J.-J., and T.S.L. designed study, data analysis, and interpretation. J.J. and K.J.-J. wrote applications for the study. J.J., K.J.-J., and T.S.L. collected data. J.J. and Z.B. performed statistical analyses. J.J. wrote a draft manuscript. K.J.-J., M.R.S., M.R.C., A.L.A.-M., A.O.G., P.B., R.S.P., C.B.P., T.C.E.-G., and T.S.L. supervised and provided valuable input into the present study. All authors revised the manuscript. All authors approved the final version.

Acknowledgments

We want to express our gratitude to our colleagues from the Department of Clinical Biochemistry in Denmark for their assistance in the laboratory data analysis for this study.

Ethics Statement

The study was approved under number 21/27006 by the Advisory Board for Clinical Quality Development Program (RKKP).

Conflicts of Interest

P.B. reports advisory board from SERB, Roche, Gilead, Abbvie. T.S.L. reports consultancy/advisory board from BMS, Novartis, Gilead, Roche;

research grant from Genentech; and travel expenses from Roche. The other authors declare no conflicts of interest.

Data Availability Statement

The datasets analyzed during the current study can be made available to the corresponding author upon reasonable request.

References

1. A. S. Ruppert, J. G. Dixon, G. Salles, et al., "International Prognostic Indices in Diffuse Large B-Cell Lymphoma: A Comparison of IPI, R-IPI, and NCCN-IPI," *Blood* 135 (2020): 2041–2048.

2. The International Non-Hodgkin's Lymphoma Prognostic Factors Project, "A Predictive Model for Aggressive Non-Hodgkin's Lymphoma," *New England Journal of Medicine* 329 (1993): 987–994.

3. L. H. Sehn, B. Berry, M. Chhanabhai, et al., "The Revised International Prognostic Index (R-IPI) Is a Better Predictor of Outcome Than the Standard IPI for Patients With Diffuse Large B-Cell Lymphoma Treated With R-CHOP," *Blood* 109 (2007): 1857–1861.

4. Z. Zhou, L. H. Sehn, A. W. Rademaker, et al., "An Enhanced International Prognostic Index (NCCN-IPI) for Patients With Diffuse Large B-Cell Lymphoma Treated in the Rituximab Era," *Blood* 123 (2014): 837–842.

5. M. Ziepert, D. Hasenclever, E. Kuhnt, et al., "Standard International Prognostic Index Remains a Valid Predictor of Outcome for Patients With Aggressive CD20+ B-Cell Lymphoma in the Rituximab Era," *Journal of Clinical Oncology* 28 (2010): 2373–2380.

6. R. Herbrecht, P. Cernohous, A. Engert, et al., "Comparison of Pixantrone-Based Regimen (CPOP-R) With Doxorubicin-Based Therapy (CHOP-R) for Treatment of Diffuse Large B-Cell Lymphoma," *Annals of Oncology* 24 (2013): 2618–2623.

7. J. F. Seymour, M. Pfreundschuh, M. Trněný, et al., "R-CHOP With or Without Bevacizumab in Patients With Previously Untreated Diffuse Large B-Cell Lymphoma: Final MAIN Study Outcomes," *Haematologica* 99 (2014): 1343–1349.

8. N. L. Bartlett, W. H. Wilson, S. H. Jung, 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," *Journal of Clinical Oncology* 37 (2019): 1790–1799.

9. A. Davies, T. E. Cummin, S. Barrans, 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 Oncology* 20 (2019): 649–662.

10. Z. Loh, R. Salvaris, G. Chong, et al., "Evolution of Eligibility Criteria for Diffuse Large B-Cell Lymphoma Randomised Controlled Trials Over 30 Years," *British Journal of Haematology* 193 (2021): 741–749.

11. M. J. Maurer, H. Ghesquières, B. K. Link, et al., "Diagnosis-To-Treatment Interval Is an Important Clinical Factor in Newly Diagnosed Diffuse Large B-Cell Lymphoma and Has Implication for Bias in Clinical Trials," *Journal of Clinical Oncology* 36 (2018): 1603–1610.

12. R. A. Harkins, S. P. Patel, M. J. Lee, et al., "Improving Eligibility Criteria for First-Line Trials for Patients With DLBCL Using a US-Based Delphi-Method Survey," *Blood Advances* 6 (2022): 2745–2756.

13. H. Tilly, F. Morschhauser, L. H. Sehn, et al., "Polatuzumab Vedotin in Previously Untreated Diffuse Large B-Cell Lymphoma," *New England Journal of Medicine* 386 (2022): 351–363.

14. U. Vitolo, M. Trneny, D. Belada, et al., "Obinutuzumab or Rituximab Plus Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone in Previously Untreated Diffuse Large B-Cell Lymphoma," *Journal of Clinical Oncology* 35 (2017): 3529–3537.

15. A. Younes, L. H. Sehn, P. Johnson, et al., "Randomized Phase III Trial of Ibrutinib and Rituximab Plus Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone in Non-Germinal Center B-Cell Diffuse Large B-Cell Lymphoma," *Journal of Clinical Oncology* 37 (2019): 1285–1295.

16. C. Récher, B. Coiffier, C. Haioun, 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 (London, England)* 378 (2011): 1858–1867.

17. S. A. Atallah-Yunes, A. Khurana, and M. Maurer, "Challenges Identifying DLBCL Patients With Poor Outcomes to Upfront Chemoimmunotherapy and Its Impact on Frontline Clinical Trials," *Leukemia & Lymphoma* 65 (2024): 430–439.

18. B. Arboe, T. C. El-Galaly, M. R. Clausen, et al., "The Danish National Lymphoma Registry: Coverage and Data Quality," *PLoS One* 11 (2016): e0157999.

19. A. S. Levey, L. A. Inker, and J. Coresh, "GFR Estimation: From Physiology to Public Health," *American Journal of Kidney Diseases* 63 (2014): 820–834.

20. B. D. Cheson, R. I. Fisher, S. F. Barrington, et al., "Recommendations for Initial Evaluation, Staging, and Response Assessment of Hodgkin and Non-Hodgkin Lymphoma: The Lugano Classification," *Journal of Clinical Oncology* 32 (2014): 3059–3067.

21. G. Tripepi, K. J. Jager, F. W. Dekker, and C. Zoccali, "Statistical Methods for the Assessment of Prognostic Biomarkers (Part I): Discrimination," *Nephrology, Dialysis, Transplantation* 25 (2010): 1399–1401.

22. H. Uno, T. Cai, M. J. Pencina, R. B. D'Agostino, and L. J. Wei, "On the C-Statistics for Evaluating Overall Adequacy of Risk Prediction Procedures With Censored Survival Data," *Statistics in Medicine* 30 (2011): 1105–1117.

23. K. P. Burnham and D. R. Anderson, "Multimodel Inference: Understanding AIC and BIC in Model Selection," *Sociological Methods & Research* 33 (2004): 261–304.

24. B. Barton and J. Peat, Test of Reliability and Agreement. *In Medical Statistics: A Guide to SPSS, Data Analysis and Critical Appraisal* (Chichester, West Sussex: BMJ Publishing Group, 2014), 317–318.

25. M. B. Davidson, A. Rice, D. Stewart, and C. Owen, "Generalizability of Landmark Clinical Trials in Diffuse Large B Cell Lymphoma to Real-World Patients: A Single-Center Retrospective Cohort Study," *Leukemia & Lymphoma* 63 (2022): 2996–2999.

26. A. Khurana, R. Mwangi, G. S. Nowakowski, et al., "Impact of Organ Function-Based Clinical Trial Eligibility Criteria in Patients With Diffuse Large B-Cell Lymphoma: Who Gets Left Behind?" *Journal of Clinical Oncology* 39 (2021): 1641–1649.

27. L. H. Sehn and G. Salles, "Diffuse Large B-Cell Lymphoma," *New England Journal of Medicine* 384 (2021): 842–858.

28. M. J. Maurer, A. Khurana, U. Farooq, et al., "The Presence of Bulky Disease and/or Very High LDH Defines a High-Risk Subset of IPI 1-2 for Eligibility in Clinical Trials of Newly Diagnosed Aggressive B-Cell Lymphoma," *Blood* 142 (2023): 4512.

29. J. Jelicic, T. S. Larsen, M. Maksimovic, and G. Trajkovic, "Available Prognostic Models for Risk Stratification of Diffuse Large B Cell Lymphoma Patients: A Systematic Review," *Critical Reviews in Oncology/Hematology* 133 (2019): 1–16.

30. A. O. Gang, M. Pedersen, F. d'Amore, et al., "A Clinically Based Prognostic Index for Diffuse Large B-Cell Lymphoma With a cut-off at 70 Years of Age Significantly Improves Prognostic Stratification: Population-Based Analysis From the Danish Lymphoma Registry," *Leukemia & Lymphoma* 56 (2015): 2556–2562.

31. J. Jelicic, K. Juul-Jensen, Z. Bukumiric, et al., "Revisiting Beta-2 Microglobulin as a Prognostic Marker in Diffuse Large B-Cell Lymphoma," *Cancer Medicine* 13 (2024): e7239.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.