ARTICLE
Molecular Diagnostics

A nomogram prognostic index for risk-stratification in diffuse large B-cell lymphoma in the rituximab era: a multi-institutional cohort study

Jun Cai[1,2], Xiaopeng Tian[1,2], Shuyun Ma[1,2], Liye Zhong[3], Wenyu Li[4], Liang Wang [5], Linlang Guo[6], Zhihua Li[7], Yudan Wu[7], Guangzheng Zhong[8], Huiqiang Huang[1,2], Zhongjun Xia[1,9], Yi Xia[1,2], Panpan Liu[1,2], Ning Su[1,2], Yu Fang[1,2], Yuchen Zhang[1,2] and Qingqing Cai[1,2]

BACKGROUND: We aimed to establish a predictive prognostic risk-stratification model for diffuse large B-cell lymphoma (DLBCL) in the rituximab era.

METHODS: The data of 1406 primary DLBCL patients from the Sun Yat-Sen University Cancer Center were analysed to establish a nomogram prognostic index (NPI) model for predicting overall survival (OS) based on pre-treatment indicators. An independent cohort of 954 DLBCL patients from three other hospitals was used for external validation.

RESULTS: Age, performance status, stage, lactate dehydrogenase, number of extranodal sites, BCL2, CD5 expression, B symptoms and absolute lymphocyte and monocyte count were the main factors of the NPI model and could stratify the patients into four distinct categories based on their predicted OS. The calibration curve demonstrated satisfactory agreement between the predicted and actual 5-year OS of the patients. The concordance index of the NPI model (0.794) was higher than the IPI (0.759) and NCCN-IPI (0.750), and similar results were obtained upon external validation. For CD5 + DLBCL patients, systemic treatment with high-dose methotrexate was associated with superior OS compared to R-CHOP-based immunochemotherapy alone.

CONCLUSIONS: We established and validated an accurate prediction model, which performed better than IPI and NCCN-IPI for prognostic stratification of DLBCL patients.

BACKGROUND

Diffuse large B-cell lymphoma (DLBCL) is one of the most common types of non-Hodgkin lymphoma (NHL), accounting for ~32% and ~38% of NHL cases in western countries and China.[1,2] The R-CHOP regimen (rituximab combination with cyclophosphamide, doxorubicin, vincristine and prednisone) is the standard treatment for all stages of DLBCL yet a considerable proportion of patients suffer from a high risk of relapse despite primary treatment. The prognosis of relapsed/refractory DLBCL patients is relatively poor, with a 5-year overall survival (OS) < 40%, compared to 60–70% for primary DLBCL patients.[3,4] Thus, it is necessary to formulate ways for the early identification of patients with a high risk of recurrence.

The International Prognostic Index (IPI) consists of five independent prognostic factors, namely: age, the Eastern Cooperative Oncology Group performance status (ECOG PS), Ann Arbor stage, lactate dehydrogenase (LDH) and the number of extranodal sites.[5] It is currently used as a guide for patients' survival prognostication at diagnosis, by classifying them into four risk groups. In the rituximab era, the revised IPI (R-IPI) and National Comprehensive Cancer Network (NCCN)-IPI models were proposed. They use similar clinical and biochemical characteristics to improve their discriminatory power of prognosis by regrouping the original IPI score or stratifying age and LDH.[6,7] However, as they failed to fully identify the extremely high-risk population,[8,9] the IPI model is still used as the benchmark for determining the prognosis of DLBCL patients.

In the past few years, significant achievements have been made in the understanding of lymphoma biology and exploring the genetic subtypes of DLBCL.[10–14] Numerous novel markers with potential prognostic significance have also been identified. Laboratory hallmarks indicators, such as absolute lymphocyte

---

[1]State Key Laboratory of Oncology in South China, Collaborative Innovation Center of Cancer Medicine, Sun Yat-sen University Cancer Center, Guangzhou, P. R. China; [2]Department of Medical Oncology, Sun Yat-sen University Cancer Center, Guangzhou, P. R. China; [3]Department of Hematology, Guangdong General Hospital, Guangdong Academy of Medical Sciences, Guangzhou, P. R. China; [4]Lymphoma Division, Cancer Center of Guangdong General Hospital, Guangdong Academy of Medical Sciences, Guangzhou, P. R. China; [5]Department of Hematology, Beijing Tongren Hospital, Capital Medical University, Beijing, P. R. China; [6]Department of Pathology, Zhujiang Hospital, Southern Medical University, Guangzhou, P. R. China; [7]Department of Oncology, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, P. R. China; [8]Department of Urology, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, P. R. China and [9]Department of Hematology Oncology, Sun Yat-sen University Cancer Center, Guangzhou, P. R. China
Correspondence: Qingqing Cai (caiqq@sysucc.org.cn)
These authors contributed equally: Jun Cai, Xiaopeng Tian, Shuyun Ma, Liye Zhong, Wenyu Li

A nomogram prognostic index for risk-stratification in diffuse large...
J Cai et al.

403

count (ALC), absolute monocyte count (AMC)and histopathologic characters, i.e. BCL2, MYC, CD5 expression and cell of origin, were applied to develop new models with improved discrimination power of prognosis.[15–25] However, they still could neither accurately predict OS or progression-free survival (PFS) nor had a substantial impact on the treatment of DLBCL patients.[26,27] Considering that there is currently no effective way to combine laboratory biomarkers and histopathologic features into the current IPI score for OS prediction, the establishment of a more robust and comprehensive prognostic nomogram incorporating clinic histopathologic variables for DLBCL patients is imminent.

Nomograms are prognostic models for predicting the outcomes of individuals and have been increasingly used in oncology.[28–30] However, their predictive value in DLBCL has not yet been adequately explored. In this study, we aimed to develop and validate an easy-to-use nomogram prognostic index (NPI) by combining IPI, histopathologic and laboratory factors with prognostic significance in DLBCL patients. We also aimed to determine whether the NPI model could more accurately predict OS compared to the IPI and NCCN-IPI models.

METHODS
Patients and treatments
We retrospectively collected the clinical data of primary DLBCL adult patients (age ≥18 years) who underwent first-line R-CHOP-based (rituximab plus other standard anthracycline-based chemotherapy) immunochemotherapy regimens. Prophylactic interventions for central nervous system (CNS) relapse were allowed for the following individuals: patients in the "high-risk" group according to CNS-IPI, CD5 + DLBCL patients and extranodal DLBCL like breast, uterus, paranasal sinuses, epidural, bone and bone marrow involvement. Patients were excluded if they were diagnosed with primary CNS lymphoma or "double-/triple-hit" lymphomas.[31] Detailed inclusion and exclusion criteria are described in the Supplementary Methods. Four institutions participated in this study, namely (1) the Sun Yat-sen University Cancer Center (SYSUCC; Guangzhou, China), (2) Guangdong General Hospital (Guangzhou, China), (3) Sun Yat-sen Memorial Hospital of Sun Yat-sen University (Guangzhou, China) and (4) Zhujiang Hospital of Southern Medical University (Guangzhou, China). Cases from SYSUCC were used as the training cohort while data from the other three hospitals were used for external validation.

Clinicopathologic factors assessments and endpoints
Baseline clinical characteristics included age, gender, height, weight, B symptoms (presence of at least one of the followings: night sweat, weight loss>10% over six months and recurrent fever >38.3 °C), ECOG PS, stage (Ann Arbor stage I-IV, performed by ^{18}F-fluorodeoxyglucose [^{18}FDG] positron emission tomography [PET] plus computed tomography according to the Lugano Classification[32]), and the number of extranodal sites. The body mass index (BMI) was calculated by dividing the body weight (measured in kilograms) with the square of height (measured in metres). Low BMI was defined as <25.0 kg/m^2.[33] LDH was considered elevated if its level was >245 U/L in the plasma.[5] A decreased absolute lymphocyte count (ALC) was defined as ≤1000/µl, and an elevated absolute monocyte count (AMC) was defined as ≥630/µL in whole blood.[19] Plasma uric acid (UA) ≥ 75th percentile was considered elevated.[8] C-reactive protein (CRP) and β-2 microglobulin (β2M) were considered elevated if their serum levels were >8 mg/L and >3.2 mg/L, respectively.[17,34] Pre-treatment biopsy samples, i.e. lymph node biopsy, were collected for immunohistochemical (IHC) analysis, providing expression positivity for BCL2 (BCL2 +: protein expression ≥50%), MYC (MYC +: protein expression ≥40%), CD20 and CD5 (CD20 +, CD5 +: expression of protein in at least a small population of the neoplastic cells),[21,35] and cell of

origin (GCB versus non-GCB) according to the Hans algorithm.[36] The expression of MYC was not included in this study due to insufficient data available. The primary endpoint was 5-year OS, measured from the start of diagnosis to the date of last follow-up or death from any cause.

Construction and validation of the NPI model
In the training cohort, risk factors selected for univariate analyses were based on previous studies and were routinely available in clinical practice. Performance of the NPI model was assessed using the concordance index (C-index), the area under the curve (AUC) of the receiver operating characteristic (ROC) curve and integrated Brier score (IBS). Calibration curves were derived based on regression analyses to determine whether the predicted probability was consistent with the actual survival of the patients. Comparisons of the predictive ability between the NPI model, IPI and NCCN-IPI were investigated by their C-indices, AUC and IBS. In the validation cohort, applying the total scores as an independent factor, each patient was stratified into risk groups according to the NPI model, and the prediction accuracy of the NPI model was also evaluated by ROC curve analysis, C-index and IBS.

Statistical analyses
Univariate analysis was used to identify potential prognostic variables correlated to OS. Significant variables ($P < 0.1$) in the univariate analyses were selected for multivariable analyses using the multivariable Cox proportional hazards model. The nomogram was established with coefficients of the independent prognostic factors weighted by the multivariate analysis. Survival curves were compared using the log-rank test. Bootstraps with 1000 resamples were used to calculate C-index. Statistical analyses to identify independent prognostic variables were performed in the SPSS 25.0 for Windows, and the NPI model was formulated using the Hmisc, rms and survival ROC packages in R, version 3.6.1 (http://www.r-project.org/). All statistical tests were two-sided, and P value <0.05 was statistically significant.

RESULTS
Patient characteristics
The entire cohort comprised 3686 primary DLBCL patients who were diagnosed between January 1, 2005 and December 30, 2015. 1326 patients were excluded from the analysis. A total of 2360 primary patients with DLBCL met all eligibility criteria and was included in this study. The training and the validation cohort consisted of 1406 and 954 patients (Supplementary Fig. 1). The baseline characteristics of DLBCL patients in each cohort are listed in Table 1. The median follow-up duration of the training and validation cohort was 81.3 (IQR 62.4–103.7) and 78.5 (IQR 65.3–92.7) months, respectively. In total, 303 (21.6%) deaths in the training cohort and 218 (22.9%) deaths in the validation cohort were recorded, and the 5-year OS were 80.8% and 78.6%, respectively.

Independent prognostic factors for OS in the training set
The results of univariate analyses are listed in Table 2. The five variables in IPI (age, LDH, ECOG PS, Ann Arbor stage and extranodal sites) were predictive factors significantly affecting OS ($P < 0.001$). Non-GCB type, BCL2 + and CD5 +, B symptoms, increased AMC, CRP and β2M and decreased ALC were also found to be associated with OS. All significant factors ($P < 0.1$) associated with survival in univariate analyses were included for multivariate analysis. Apart from the five indicators of the IPI model, BCL2 + (HR 1.663, 95% CI 1.256–2.203, $P < 0.001$), CD5 + (HR 2.072, 95% CI 1.565–2.742, $P < 0.001$), B symptoms (HR 1.338, 95% CI 1.043–1.717, $P = 0.022$), ALC (HR 1.430, 95% CI 1.093–1.872, $P = 0.009$), and AMC (HR 1.867, 95% CI 1.463–2.381, $P < 0.001$) were also identified as independent

A nomogram prognostic index for risk-stratification in diffuse large…
J Cai et al.

404

Table 1. Clinical characteristics of DLBCL patients in the training and validation cohorts.

Characteristic	Training cohort					Validation cohort					P*
	Patients (n = 1406)	Low-risk (n = 773)	Intermediate-risk (n = 319)	High-risk (n = 217)	Extremely high-risk (n = 97)	Patients (n = 954)	Low-risk (n = 502)	Intermediate-risk (n = 234)	High-risk (n = 130)	Extremely high-risk (n = 88)	
Sex, n (%)											
Male	794 (56.5)	432 (55.9)	181 (56.7)	127 (58.5)	54 (55.7)	561 (58.8)	298 (59.4)	137 (58.5)	76 (58.5)	50 (56.8)	0.261
Female	612 (43.5)	341 (44.1)	138 (43.3)	90 (41.5)	43 (44.3)	393 (41.2)	204 (40.6)	97 (41.5)	54 (41.5)	38 (43.2)	
Age (years), n (%)											
≤60	984 (70.0)	661 (85.8)	160 (50.2)	132 (60.8)	31 (32.0)	641 (67.2)	419 (83.5)	121 (51.7)	77 (59.2)	24 (27.3)	0.150
>60	422 (30.0)	112 (14.5)	159 (49.8)	85 (39.2)	66 (68.0)	313 (32.8)	83 (16.5)	113 (48.3)	53 (40.8)	64 (72.7)	
LDH (U/L), n (%)											
≤ULN	852 (60.6)	638 (82.5)	163 (51.1)	44 (20.3)	7 (7.2)	577 (60.5)	414 (82.5)	122 (52.1)	36 (27.7)	5 (5.7)	0.955
>ULN	554 (39.4)	135 (17.5)	156 (48.9)	173 (79.7)	90 (92.8)	377 (39.5)	88 (17.5)	112 (47.9)	94 (72.3)	83 (94.3)	
ECOG PS, n (%)											
<2	1313 (93.4)	762 (98.6)	308 (96.6)	188 (86.6)	55 (56.7)	889 (93.2)	497 (99.0)	226 (96.6)	109 (83.8)	57 (64.8)	0.850
≥2	93 (6.6)	11 (1.4)	11 (3.4)	29 (13.4)	42 (43.3)	65 (6.8)	5 (1.0)	8 (3.4)	21 (16.2)	31 (35.2)	
Ann Arbor stage, n (%)											
I–II	774 (55.0)	620 (80.2)	126 (39.5)	27 (12.4)	1 (1.0)	498 (52.2)	370 (73.7)	98 (41.9)	23 (17.7)	7 (8.0)	0.173
III–IV	632 (45.0)	153 (19.8)	193 (60.5)	190 (87.6)	96 (99.0)	456 (47.8)	132 (26.3)	136 (58.1)	107 (82.3)	81 (92.0)	
No. of extranodal sites, n (%)											
<2	1179 (83.9)	758 (98.1)	275 (86.2)	115 (53.0)	31 (32.0)	779 (81.7)	482 (96.0)	193 (82.5)	76 (58.5)	28 (31.8)	0.163
≥2	227 (16.1)	15 (1.9)	44 (13.8)	102 (47.0)	66 (68.0)	175 (18.3)	20 (4.0)	41 (17.5)	54 (41.5)	60 (68.2)	
BMI (kg/m²) n (%)											
<25	1139 (81.0)	614 (79.4)	269 (84.3)	174 (80.2)	82 (84.5)	784 (82.2)	411 (81.9)	196 (83.8)	105 (80.8)	72 (81.8)	0.473
≥25	267 (19.0)	159 (20.6)	50 (15.7)	43 (19.8)	15 (15.5)	170 (17.8)	91 (18.1)	38 (16.2)	25 (19.2)	16 (18.2)	
BCL2, n (%)											
Positive	958 (68.1)	464 (60.0)	235 (73.7)	172 (79.3)	87 (89.7)	634 (66.5)	268 (53.4)	173 (73.9)	110 (84.6)	82 (94.3)	0.393
Negative	448 (31.9)	309 (40.0)	84 (26.3)	45 (20.7)	10 (10.3)	320 (33.5)	234 (46.6)	61 (26.1)	20 (15.4)	5 (5.7)	
CD5, n (%)											
Positive	171 (12.2)	35 (4.5)	59 (18.5)	46 (21.2)	31 (32.0)	137 (14.4)	28 (5.6)	41 (17.5)	33 (25.4)	35 (39.8)	0.120
Negative	1235 (87.8)	738 (95.5)	260 (81.5)	171 (78.8)	66 (68.0)	817 (85.6)	474 (94.4)	193 (82.5)	97 (74.6)	53 (60.2)	
Coo by IHC, n (%)											
GCB type	633 (45.0)	383 (49.5)	134 (42.0)	79 (36.4)	37 (38.1)	428 (44.9)	241 (48.0)	95 (40.6)	58 (44.6)	34 (38.6)	0.940
Non-GCB type	773 (55.0)	390 (50.5)	185 (58.0)	138 (63.6)	60 (61.9)	526 (55.1)	261 (52.0)	139 (59.4)	72 (55.4)	54 (61.4)	
B symptoms, n (%)											
Present	320 (22.8)	88 (11.4)	82 (25.7)	92 (42.4)	58 (59.8)	203 (21.3)	43 (8.6)	47 (20.1)	60 (46.2)	53 (60.2)	0.395
Absent	1086 (77.2)	685 (88.6)	237 (74.3)	125 (57.6)	39 (40.2)	751 (78.7)	459 (91.4)	187 (79.9)	70 (53.8)	35 (39.8)	
ALC (/μL), n (%)											
≤1000	243 (17.3)	63 (8.2)	55 (17.2)	83 (38.2)	42 (43.3)	186 (19.5)	33 (6.6)	56 (23.9)	55 (42.3)	42 (47.7)	0.171
>1000	1163 (82.7)	710 (91.8)	264 (82.8)	134 (61.8)	55 (56.7)	768 (80.5)	469 (93.4)	178 (76.1)	75 (57.7)	46 (52.3)	
AMC (/μL), n (%)											
<630	932 (66.3)	644 (83.3)	178 (55.8)	83 (38.2)	27 (27.8)	597 (62.6)	401 (79.9)	117 (50.0)	50 (38.5)	29 (33.0)	0.064
≥630	474 (33.7)	129 (16.7)	141 (44.2)	134 (61.8)	70 (72.2)	357 (37.4)	101 (20.1)	117 (50.0)	80 (61.5)	59 (67.9)	

A nomogram prognostic index for risk-stratification in diffuse large...
J Cai et al.

405

Table 1. continued

Characteristic	Training cohort					Validation cohort					P*
	Patients (n = 1406)	Low-risk (n = 773)	Intermediate-risk (n = 319)	High-risk (n = 217)	Extremely high-risk (n = 97)	Patients (n = 954)	Low-risk (n = 502)	Intermediate-risk (n = 234)	High-risk (n = 130)	Extremely high-risk (n = 88)	
CRP (mg/L), n (%)											0.358
≤8	852 (60.6)	584 (75.5)	163 (51.1)	76 (35.0)	29 (29.9)	596 (62.5)	375 (74.7)	126 (53.8)	59 (45.4)	36 (40.9)	
>8	554 (39.4)	189 (24.5)	156 (48.9)	141 (65.0)	68 (70.1)	358 (37.5)	127 (25.3)	108 (46.2)	71 (54.6)	52 (59.1)	
β2M (mg/L), n (%)											0.097
≤3.2	823 (58.5)	522 (67.5)	156 (48.9)	105 (48.4)	40 (41.2)	591 (61.9)	365 (72.7)	142 (60.7)	51 (39.2)	33 (37.5)	
>3.2	583 (41.5)	251 (32.5)	163 (51.1)	112 (51.6)	57 (58.8)	363 (38.1)	137 (27.3)	92 (39.3)	79 (60.8)	55 (62.5)	
UA (umol/L), n (%)											0.993
<75th percentile	1054 (75.0)	600 (77.6)	236 (74.0)	157 (72.4)	61 (62.9)	715 (74.9)	388 (77.3)	172 (73.5)	94 (72.3)	61 (69.3)	
≥75th percentile	352 (25.0)	173 (22.4)	83 (26.0)	60 (27.6)	36 (37.1)	239 (25.1)	114 (22.7)	62 (26.5)	36 (27.7)	27 (30.7)	
IPI, n (%)											0.256
Low (0–1)	830 (59.0)	699 (90.4)	117 (36.7)	14 (6.5)	0 (0.0)	537 (56.3)	437 (87.1)	88 (37.6)	12 (9.2)	0 (0.0)	
Low intermediate (2)	285 (20.3)	74 (9.6)	146 (45.8)	62 (28.6)	3 (3.1)	223 (23.4)	62 (12.4)	109 (46.6)	46 (35.4)	6 (6.8)	
High intermediate (3)	210 (14.9)	0 (0.0)	54 (16.9)	123 (56.7)	33 (34.0)	133 (13.9)	3 (0.6)	37 (15.8)	63 (48.5)	30 (34.1)	
High (≥4)	81 (5.8)	0 (0.0)	2 (0.6)	18 (8.3)	61 (62.9)	61 (6.4)	0 (0.0)	0 (0.0)	9 (6.9)	52 (59.1)	
Treatment, n (%)											0.067
R-CHOP	1261 (89.7)	700 (90.4)	281 (88.1)	193 (88.9)	87 (89.7)	877 (91.9)	458 (91.2)	216 (92.3)	120 (92.3)	83 (94.3)	
R-CHOP-like	145 (10.3)	74 (9.6)	38 (11.9)	24 (11.1)	10 (10.3)	77 (8.1)	44 (8.8)	18 (7.7)	10 (7.7)	5 (5.7)	
Receiving HD-MTX, n (%)											0.217
Yes	115 (8.2)	47 (6.1)	31 (9.7)	30 (13.8)	7 (7.2)	92 (9.6)	46 (9.2)	22 (9.4)	17 (13.7)	7 (8.0)	
No	1291 (91.8)	726 (93.9)	288 (90.3)	187 (86.2)	90 (92.3)	862 (90.4)	456 (90.8)	212 (90.6)	113 (86.9)	81 (92.0)	
Response to the treatment, n (%)											0.750
CR	947 (67.4)	586 (75.8)	207 (64.9)	106 (48.8)	48 (49.5)	635 (66.6)	374 (74.5)	159 (67.9)	64 (49.2)	38 (43.2)	
PR	362 (25.7)	149 (19.3)	92 (28.8)	88 (40.6)	33 (34.0)	261 (27.4)	110 (21.9)	64 (27.4)	52 (40.0)	35 (39.8)	
SD	58 (4.1)	23 (3.0)	15 (4.7)	13 (6.0)	7 (7.2)	34 (3.6)	11 (2.2)	5 (2.1)	9 (6.9)	9 (10.2)	
PD	39 (2.8)	15 (1.9)	5 (1.6)	10 (4.6)	9 (9.3)	24 (2.5)	7 (1.4)	6 (2.6)	5 (3.8)	6 (6.8)	

ECOG PS Eastern Cooperative Oncology Group Performance Status, LDH lactate dehydrogenase (ULN: 245 U/L), ULN upper limit of normal, BMI body mass index, ALC absolute lymphocyte count, AMC absolute monocyte count, CRP C-reactive protein, UA uric acid, β2M β-2 microglobulin (ULN: 2.52 mg/L), IPI International Prognostic Index, R-CHOP rituximab, cyclophosphamide, doxorubicin or liposomal doxorubicin, vincristine, prednisone, R-CHOP-like included dose-adjusted EPOCH (etoposide, prednisone, vincristine, cyclophosphamide, doxorubicin) plus rituximab, RCEOP (rituximab, cyclophosphamide, epirubicin, vincristine, prednisone) et al., HD-MTX high-dose methotrexate, CR complete remission, PR partial remission, SD stable disease, PD progressive disease.
*P value was calculated by Pearson Chi-Square test between patients in the training and validation cohorts.

A nomogram prognostic index for risk-stratification in diffuse large…
J Cai et al.

406

Table 2. Univariate and multivariate analysis of the training and validation cohorts.

Variable	Univariate				Multivariate			
	Training cohort ($n = 1406$)		Validation cohort ($n = 954$)		Training cohort ($n = 1406$)		Validation cohort ($n = 954$)	
	Hazard ratio (95% CI)	P	Hazard ratio (95% CI)	P	Hazard ratio (95% CI)	P	Hazard ratio (95% CI)	P
Sex (male)	1.112 (0.884–1.398)	0.364	1.050 (0.800–1.377)	0.727				
Age >60 (years)	2.478 (1.978–3.106)	<0.001	2.065 (1.582–2.696)	<0.001	2.380 (1.893–2.992)	<0.001	1.770 (1.339–2.340)	<0.001
LDH > ULN (U/L)	3.223 (2.549–4.074)	<0.001	3.183 (2.412–4.199)	<0.001	1.576 (1.197–2.077)	0.001	1.633 (1.184–2.252)	0.003
ECOG PS ≥ 2	3.952 (2.908–5.371)	<0.001	4.701 (3.340–6.616)	<0.001	2.754 (2.008–3.776)	<0.001	2.165 (1.487–3.153)	<0.001
Ann Arbor stage, III–IV	3.638 (2.835–4.667)	<0.001	2.692 (2.022–3.584)	<0.001	1.693 (1.257–2.280)	0.001	1.364 (0.965–1.929)	0.079
No. of extranodal sites ≥2	3.941 (3.115–4.986)	<0.001	3.479 (2.643–4.581)	<0.001	1.755 (1.336–2.306)	<0.001	1.779 (1.282–2.468)	0.001
BMI < 25 kg/m^2	1.113 (0.827–1.497)	0.480	1.052 (0.738–1.499)	0.779				
BCL2 +	1.815 (1.382–2.385)	<0.001	2.599 (1.831–3.689)	<0.001	1.663 (1.256–2.203)	<0.001	1.939 (1.352–2.782)	<0.001
CD5 +	2.302 (1.753–3.025)	<0.001	2.489 (1.842–3.362)	<0.001	2.072 (1.565–2.742)	<0.001	1.587 (1.154–2.183)	0.005
COO (GCB subtype)	0.736 (0.584–0.927)	0.009	0.948 (0.725–1.240)	0.696	0.921 (0.728–1.165)	0.491		
B symptoms	2.084 (1.645–2.639)	<0.001	2.465 (1.867–3.255)	<0.001	1.338 (1.043–1.717)	0.022	1.392 (1.023–1.894)	0.035
ALC ≤ 1000/μL	2.164 (1.686–2.778)	<0.001	2.614 (1.975–3.460)	<0.001	1.430 (1.093–1.872)	0.009	1.608 (1.199–2.158)	0.002
AMC ≥ 630/μL	2.710 (2.162–3.399)	<0.001	2.229 (1.707–2.913)	<0.001	1.867 (1.463–2.381)	<0.001	1.707 (1.298–2.245)	<0.001
CRP > 8 mg/L	2.166 (1.726–2.716)	<0.001	1.763 (1.351–2.302)	<0.001	1.128 (0.880–1.447)	0.342	1.002 (0.751–1.337)	0.988
β2GM > 3.2 mg/L	1.546 (1.234–1.937)	<0.001	1.589 (1.216–2.076)	0.001	1.030 (0.817–1.298)	0.804	0.991 (0.747–1.314)	0.950
UA ≥ 75th percentile	1.061 (0.820–1.373)	0.652	0.997 (0.733–1.357)	0.984				

LDH lactate dehydrogenase (ULN: 245 U/L), *ULN* upper limit of normal, *ECOG PS* Eastern Cooperative Oncology Group Performance Status, *COO* cell of origin, *BMI* body mass index, *ALC* absolute lymphocyte count, *AMC* absolute monocyte count, *UA* uric acid, *CRP* C-reactive protein, *β2M* β-2 microglobulin.

prognostic factors for OS (Table 2); unlike non-GCB type, CRP and β2M ($P > 0.05$).

Development of the NPI model

The risk-stratification NPI model (Fig. 1a) was established by incorporating the ten independent prognostic factors from multivariate analysis. ECOG PS was the most significant factor for OS among all the parameters, followed by age and CD5 expression, whereas, B symptoms and ALC had a moderate impact on survival. Based on the NPI model, the estimated survival probability could be easily determined by adding the total score of the ten prognostic factors. The calibration curves of the NPI model for 5-year OS prediction demonstrated promising agreement between the predicted and actual outcome (Fig. 1b).

After sorting the NPI model based on total score, four risk groups with distinct OS were identified in the training cohort. They were classified as the low-risk (points: 0–186.4; 55.0%), intermediate-risk (points: 186.4–279.4; 22.7%), high-risk (points: 279.4–377.2; 15.4%) and extremely high-risk (points: >377.2; 6.9%) groups, which corresponded to the 5-year OS rate of 94.3% (95% CI 93.5%–95.1%, $P < 0.001$), 79.7% (95% CI 77.4%–82.0%, $P < 0.001$), 56.9% (95% CI 53.6%–60.2%, $P < 0.001$) and 30.4% (95% CI 25.8%–35.0%, $P < 0.001$) (Fig. 2a). The 5-year OS prediction accuracy determined by the C-index was 0.794 (95% CI 0.770–0.817) in the training cohort. Additional analysis showed that the NPI model had higher prognostic accuracy for OS prediction than the ten independent indicators alone (Fig. 2c and Supplementary Figs. 2 and 3).

Validation of the NPI model

Using the external validation cohort, good agreement between the predicted and actual 5-year OS of the NPI model was observed (Fig. 1c). In total, 52.6% of the patients from the validation cohort were classified as low-risk, 24.5% as intermediate-risk, 13.6% as high-risk and 9.2% as extremely high-risk group, with 5-year OS rate corresponding to 93.8% (95% CI 92.7–94.9%, $P < 0.001$), 77.6%

(95% CI 75.0–80.2%, $P < 0.001$), 53.4% (95% CI 49.1–57.7%, $P < 0.001$), and 29.6% (95% CI 24.6–34.6%, $P < 0.001$), respectively (Fig. 2b). The C-index of the NPI model for predicting the 5-year OS in the validation cohort was 0.791 (95% CI 0.764–0.817). Similarly, upon external validation, the NPI model demonstrated higher prognostic accuracy for OS prediction than the ten independent indicators alone (Fig. 2d and Supplementary Figs. 4 and 5).

Comparison of OS between models

The predictive power of the 5-year OS prediction of the NPI model was compared to the current IPI and NCCN-IPI risk-stratification models. The NPI model displayed better predictive accuracy for OS prediction than the IPI and NCCN-IPI in both the training (Supplementary Fig. 6) and validation cohorts (Supplementary Fig. 7). ROC analysis showed that the NPI model had higher prognostic accuracy for OS than the IPI (AUC in the training cohort: 0.829 versus 0.785, $P = 0.035$; AUC in the validation cohort: 0.830 versus 0.768, $P = 0.014$) and NCCN-IPI (AUC in the training cohort: 0.829 versus 0.770, $P = 0.004$; AUC in the validation cohort: 0.830 versus 0.744, $P < 0.001$) (Fig. 3a, b).

For the training cohort, the C-index for the NPI model was 0.794 (95% CI 0.770–0.817), higher than that for the IPI (0.759; 95% CI, 0.734–0.784, $P < 0.001$) and NCCN-IPI (0.750; 95% CI, 0.724–0.776, $P < 0.001$). In the validation cohort, the C-index for the NPI model prediction (0.791, 95% CI 0.764–0.817) was also greater than for the IPI (0.738; 95% CI, 0.705–0.770, $P < 0.001$) and NCCN-IPI (0.741; 95% CI, 0.710–0.772, $P < 0.001$) (Fig. 3c). IBS of the NPI for OS prediction demonstrated better performance than the IPI and NCCN-IPI (Fig. 3c). These results indicated that the NPI model could be a useful predictive model with the good discriminative ability and clinical utility for DLBCL patients in the rituximab era.

R-CHOP-based immunochemotherapy plus systemic high-dose methotrexate (HD-MTX) improves survival in CD5 + patients

In the training cohort, 1261 (89.7%) patients received R-CHOP therapy, and 115 (8.2%) patients received CNS prophylaxis using

A nomogram prognostic index for risk-stratification in diffuse large...
J Cai et al.

407

a

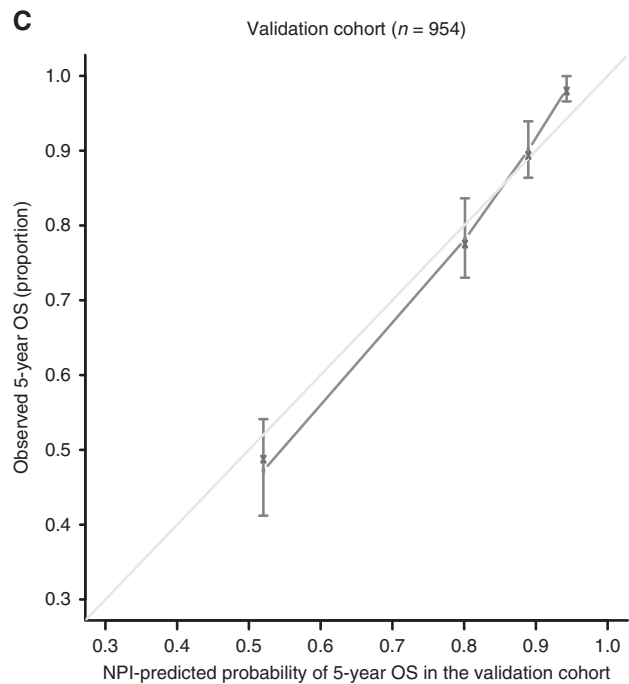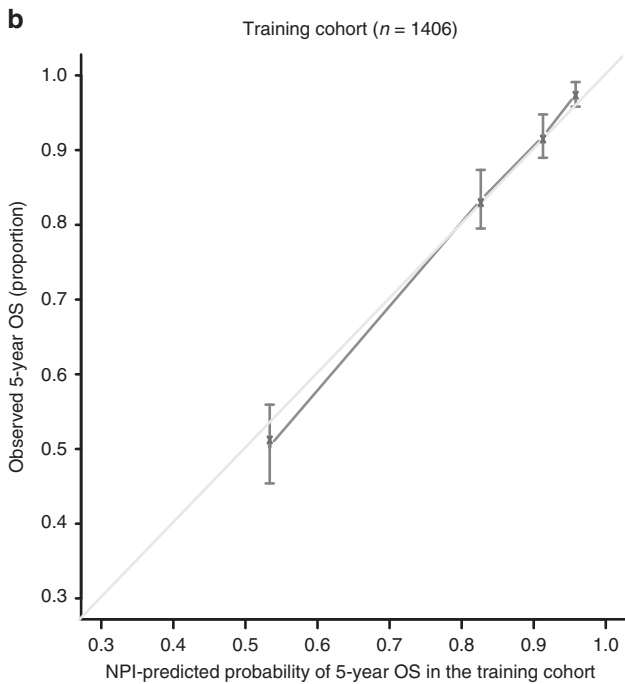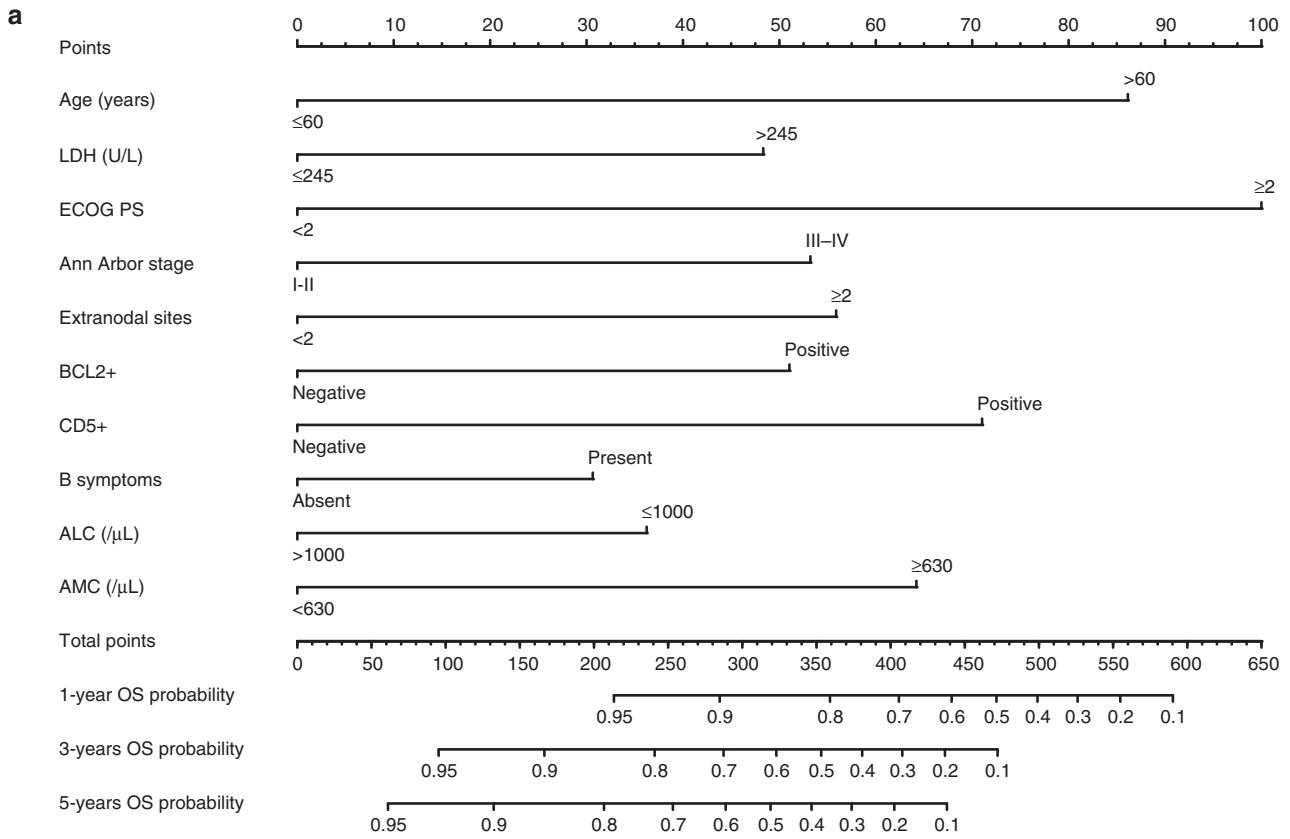**b** Training cohort (n = 1406)

c Validation cohort (n = 954)

Fig. 1 Construction of the NPI model and calibration curve of NPI model. a The NPI model included age, ECOG PS, LDH, Ann Arbor stage, extranodal sites, B symptoms, ALC, AMC, BCL2 and CD5. ECOG PS Eastern Cooperative Oncology Group Performance Status, LDH lactate dehydrogenase, ALC absolute lymphocyte count, AMC absolute monocyte count. **b**, **c** Calibration curve for predicting overall survival (OS) at 5 years in the training (n = 1406) and the validation cohort (n = 954). OS is plotted on the y-axis; NPI model-predicted probability of 5-year OS is plotted on the x-axis.

A nomogram prognostic index for risk-stratification in diffuse large…
J Cai et al.

408
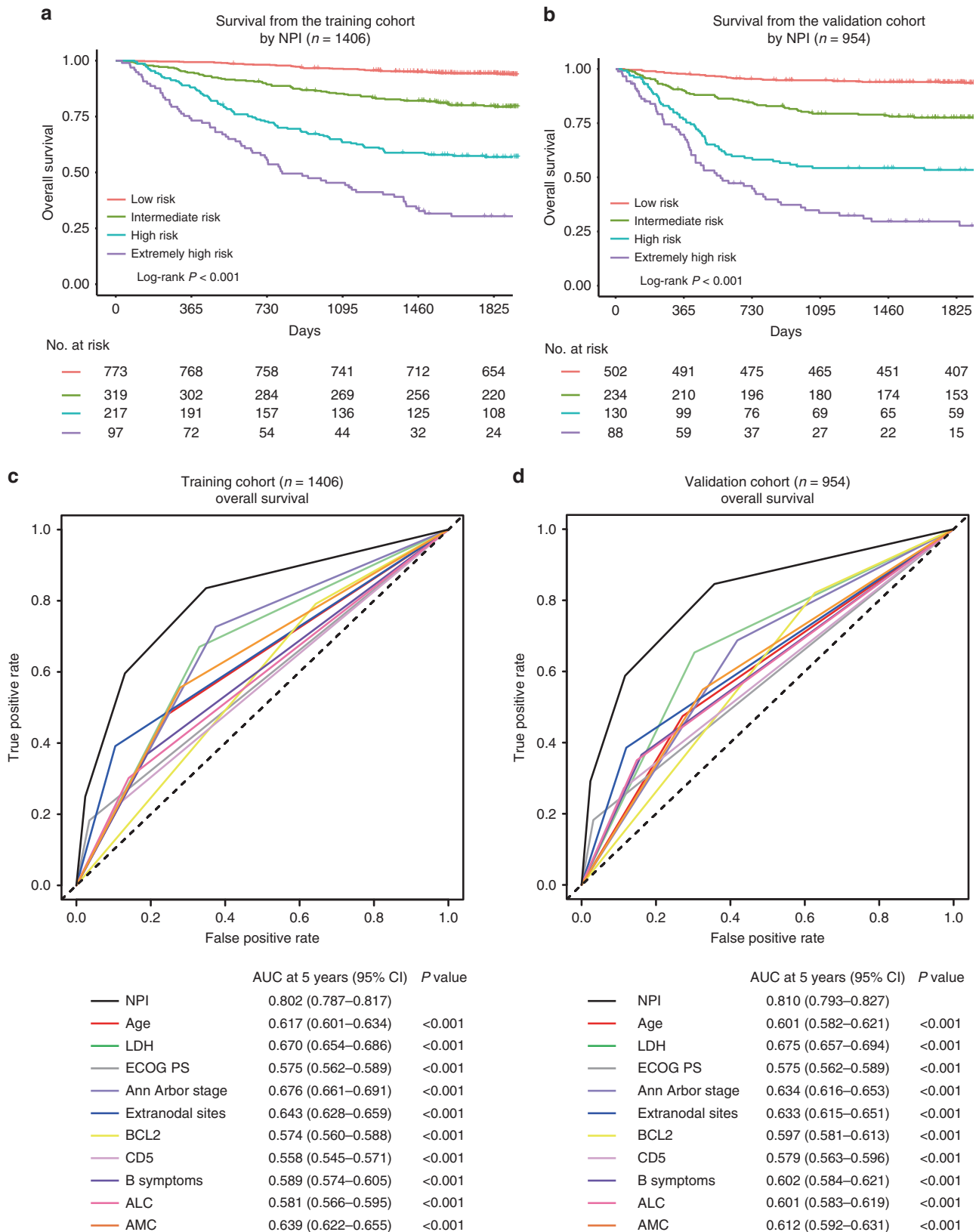
Fig. 2 NPI model to predict overall survival in DLBCL patients. a, b Subgroups of patients with different NPI model scores (low-risk: 0–184.6; intermediate-risk: 186.4–279.4; high-risk: 279.4–377.2 and extremely high-risk: >377.2) showed distinct overall survival (OS) in the training (*n* = 1406) and the validation cohort (*n* = 954). **c, d** Receiver operating characteristic (ROC) curves and the area under the curve (AUCs) at 5 years to assess the prediction performance of NPI compared with the ten clinicopathological risk indicators in the training (*n* = 1406) and the validation cohort (*n* = 954).

A nomogram prognostic index for risk-stratification in diffuse large…
J Cai et al.

409

a

Training cohort (n = 1406)
overall survival

b

Validation cohort (n = 954)
overall survival

	AUC at 5 years (95% CI)	P value
NPI	0.829 (0.844–0.815)	
IPI	0.785 (0.800–0.770)	0.035
NCCN-IPI	0.770 (0.784–0.755)	0.004

	AUC at 5 years (95% CI)	P value
NPI	0.830 (0.857–0.814)	
IPI	0.768 (0.787–0.750)	0.014
NCCN-IPI	0.744 (0.763–0.725)	<0.001

c

	Training cohort (n = 1406)			Validation cohort (n = 954)		
	C-indexes (95% CI)	P	IBS	C-indexes (95% CI)	P	IBS
NPI	0.794 (0.770–0.817)		0.078	0.791 (1.764–0.817)		0.101
IPI	0.518 (0.734–0.784)	<0.001	0.083	0.738 (0.705–0.770)	<0.001	0.107
NCCN-IPI	0.500 (0.724–0.776)	<0.001	0.085	0.741 (0.710–0.772)	<0.001	0.110

Fig. 3 Comparison of the predictive performance between NPI, IPI and NCCN-IPI models. a, b ROC curves and AUCs for predicting 5-year overall survival (OS) of NPI model, IPI and NCCN-IPI in the training (n = 1406) and the validation cohort (n = 954). **c** C-indexes and IBS of NPI, IPI and NCCN-IPI models in the training (n = 1406) and the validation cohort (n = 954).

HD-MTX. In the validation cohort, 877 (91.9%) patients received R-CHOP therapy, and 92 (9.6%) patients received HD-MTX (Table 1 and Supplementary Tables 1 and 2). Based on the regimens made no significant difference in the 5-year OS, in either the training cohort (81.1% versus 80.2%, $P = 0.993$; extremely high-risk group: 32.2% versus 16.7%, $P = 0.981$; Supplementary Fig. 8a, c) or the validation cohort (78.5% versus 82.7%, $P = 0.454$; extremely high-risk group: 29.7% versus 28.6%, $P = 0.796$; Supplementary Fig. 8b, d) between the R-CHOP or R-EPOCH (rituximab, etoposide, prednisone, vincristine, cyclophosphamide and doxorubicin) immunochemotherapy. Systemic HD-MTX did not provide additional benefit in either the whole cohort (training cohort: 5-year OS: 78.0% versus 81.1%, $P = 0.428$; validation cohort: 5-year OS: % 78.6 versus 78.7%, $P = 819$; Supplementary Fig. 8e, f) or the extremely high-risk group (training cohort: 5-year OS: 71.4% versus 27.6%, $P = 0.088$; validation cohort: 5-year OS: 50.0% versus 27.8%%, $P = 0.079$; Supplementary Fig. 8g, h).

We found a difference in the CNS relapse rates between CD5 + and CD5- DLBCL. In the training cohort, 52 (3.7%) patients experienced CNS relapse. The CNS-relapse rates were 1.6% and 2.1% for CD5 + and CD5- DLBCL, respectively. In the validation cohort, 38 (4.0%) patients experienced CNS relapse. The CNS relapse rates were 1.7% and 2.3% for CD5 + and CD5- DLBCL, respectively.

Comparison of baseline characteristics between CD5 + and CD5- DLBCL is shown in Supplementary Table 3. In the training cohort, ECOG PS ≥ 2, BCL2 +, GCB type, AMC ≥ 630/μL, CNS involvement, and received HD-MTX were more frequent in CD5 + DLBCL than those in the CD5- DLBCL. While, in the validation cohort, ECOG PS ≥ 2, more than one extranodal site involvement, BCL2 +, B symptoms, ALC ≤ 1000/μL, and higher IPI score were more frequent in CD5 + DLBCL patients.

To elucidate the effect of HD-MTX on survival outcomes for CD5 + and CD5- DLBCL patients, we further analysed the CNS relapse rates for these patients that received R-CHOP-based immunochemotherapy with or without HD-MTX. In the training
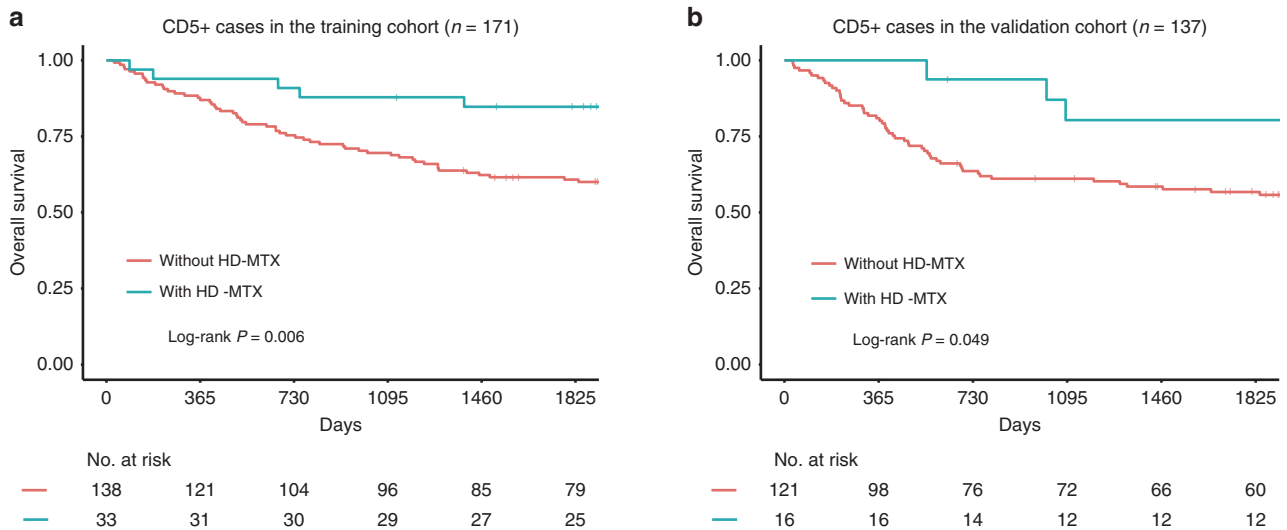
A nomogram prognostic index for risk-stratification in diffuse large…
J Cai et al.

410

Fig. 4 Kaplan–Meier survival analysis of overall survival stratified by treatment. **a**, **b** R-CHOP-based therapy alone versus R-CHOP-based therapy combined with systemic high-dose methotrexate (HD-MTX) for CD5 + patients in the training ($n = 171$) and the validation cohort ($n = 137$).

cohort, the CNS relapse rates of patients with or without HD-MTX therapy were 0.2% and 3.5%, respectively. In total, 61 (4.3%) CD5 + patients without HD-MTX therapy died, of which 21 (34.4%) cases died from CNS relapse. In the validation cohort, CNS relapse rates of patients with or without HD-MTX therapy were 0.6% and 3.4%, respectively. 56 (5.9%) CD5 + patients without HD-MTX therapy died, of which 15 (26.8%) cases died from CNS relapse. In CD5 + DLBCL patients, R-CHOP-based immunochemotherapy plus systemic HD-MTX was associated with significantly longer OS in both the training (5-year OS rate: 84.7% versus 60.8%, $P = 0.006$) and validation cohorts (5-year OS rate: 80.4% versus 56.7%, $P = 0.049$; Fig. 4a, b), compared with R-CHOP-based immunochemotherapy alone.

DISCUSSION
Based on the data from four medical institutions in China, we established and validated an NPI model that incorporated pre-treatment clinical, molecular, and pathological factors to predict the survival outcomes of DLBCL patients. The NPI model consisted of ten variables, namely age, ECOG PS, LDH, Ann Arbor stage, extranodal sites, B symptoms, ALC, AMC, BCL2 and CD5 expression, showing higher predictive performance than the IPI and NCCN-IPI models.

Decreased ALC was included in our NPI model as an independent risk factor of overall survival, and it seemed to be more important in the era of rituximab as compared to the era of CHOP chemotherapy alone.[15,19] Contrary to ALC, elevated AMC had a negative impact on the prognosis of DLBCL.[37] A low ratio of ALC:AMC was found to have the most significant prognostic value in DLBCL patients[38] in the rituximab era. However, the cut-off point of the ratio varies between studies and remains controversial,[38–40] so 1000/μL and 630/μL were chosen as the cut-off points for ALC and AMC as they have the considerable predictive ability and were more convenient to use.[19] The poor prognosis of patients with BCL2 +, MYC + and CD5 + have been reported in several studies of DLBCL.[20–23] Similar to our observations, the outcomes of BCL2 + and CD5 + patients were worse than those BCL2- and CD5- patients. Although MYC expression was prognostic for OS in DLBCL, we could not assess MYC expression because of insufficient data. Nowadays, the introduction of targeted drugs targeting specific proteins in DLBCL may gradually change the first-line treatment model for DLBCL, the inhibition of

BCL2 combined with the R-CHOP regimen demonstrated promising anti-tumour activity in patients with BCL2 + DLBCL,[41] suggesting that these molecular pathological alterations could not only be used as biomarkers for predicting prognosis but also become potential therapeutic targets.

The cell of origin (COO) measured by IHC was not significant in our study after incorporating factors into multivariate analysis. The prognostic significance of COO is controversial, as conclusions were various in different studies. Although some retrospective studies showed a survival advantage for GCB-type DLBCL distinguished by gene expression profiling (GEP),[42,43] the Hans algorithm commonly used in clinical practice only coincides with GEP for COO phenotype classification at approximately 79%.[36] Accordingly, we speculate that it may affect the prognostic value of COO to some extent.

No survival benefit was observed from R-EPOCH therapy in DLBCL patients classified as an extremely high-risk group by the NPI, this indicates that intensive chemotherapy might be of not benefit to these patients, and future studies should focus on novel targeted agents with less toxic and more efficacious. Interestingly, we found that CNS progression rates in CD5 + DLBCL patients were slightly higher than that of CD5- DLBCL in both the training and validation cohorts, and the overall survival of CD5 + patients who received HD-MTX was superior to those who didn't. Studies have shown that CD5 + DLBCL often have some distinct clinical characteristics which could be associated with worse survival and a high incidence of CNS relapse. As HD-MTX is a chemotherapeutic drug that can cross the blood-brain barrier, the combined use of standard treatment with HD-MTX may bring survival benefits for CD5 + DLBCL.[44,45] In contrast, a considerable proportion of the CD5 + patients who did not receive HD-MTX died of CNS relapse. Recently, a Phase 2 clinical study showed that CD5 + DLBCL patients could benefit from DA-EPOCH-R combined with systemic HD-MTX, which is consistent with our findings. HD-MTX is often used in the treatment of primary CNS DLBCL or as a preventive treatment after systemic therapy in high-risk patients. However, no consensus has been reached concerning the treatment of preventive interventions, and how to accurately identify patients with high-risk for CNS progression remains a challenge. The CNS-IPI Consists of six factors: five factors from IPI, and involvement of the kidney and/or adrenal glands. It has been widely used to estimate the risk of CNS recurrence/progression and guide therapeutic intervention in primary DLBCL patients in

A nomogram prognostic index for risk-stratification in diffuse large…
J Cai et al.

411

the rituximab era. According to the CNS-IPI, the high-risk population has a 10% risk of CNS recurrence, and preventive interventions should be considered.[46]

Although the NPI model did not identify patients who might benefit from HD-MTX, its prognostic predictive power cannot be ignored. Based on the proposed NPI model, ~9% of DLBCL patients were classified into the extremely high-risk group and had an extremely poor prognosis. The NPI model not only helps to predict survival outcomes but also contributes to future clinical trials design. It might be meaningful to explore novel therapies with less toxicity but more effective for these "highest risk" cases. At present, new therapies such as lenalidomide, ibrutinib and bortezomib combined with R-CHOP immunochemotherapy are undergoing clinical trials. Although no survival benefits for OS have been observed in the entire group or subgroup of those patients, the potential efficacy of those new drugs in extremely high-risk patients is not yet clear and worth exploring. In the future, prospective trials are needed to establish more effective therapies as a standard treatment for extremely high-risk DLBCL patients based on the NPI model.

Although the NPI model showed good accuracy in predicting prognosis, it still had the following limitations. First, considering clinical utility, the prognostic factors we used were limited to common clinical and histological features, and genetic markers were not included as they are not routinely available. Second, the results of IHC staining could be differ across laboratories, including differences between technicians, laboratory methods and antibody manufacturers, which could have to a certain extent affect the degree of antibody binding. Third, since the current research was mainly conducted in endemic areas of China, it remains unclear whether the NPI could be applied to patients in other regions. Lastly, this research was based on retrospective clinical data, and the predictive ability of the NPI model should be further validated in larger and prospective studies.

In summary, we developed and validated an NPI model for the risk-stratification of primary DLBCL patients and could be used as a useful tool for pre-treatment clinical evaluation of patients' survival. Although the NPI model does not contain genetic indicators such as whole exome or whole genome variables, future treatment of precision therapy will be based on the combination of the NPI model (which contains clinical, laboratory and histopathological information) and genetic indicators to guide the selection of different R-CHOP-combined therapy.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

J.C., X.P.T., S.Y.M., L.Y.Z. and W.Y.L. contributed to study design, statistical analysis and figure and tables preparation. L.Y.Z., W.Y.L., L.W., L.L.G., Z.H.L., Y.D.W., G.Z.Z., N.S., Y.F., Y.C.Z. and P.P.L. performed data collecting. J.C., X.P.T., S.Y.M. and Q.Q.C. performed manuscript writing and reviewing. Q.Q.C. conceived and designed this study that led to the submission, acquired the data and played an important role in interpreting the results, she is also responsible for all aspects of the work to sure that all questions of the work are appropriately investigated and resolved. All authors agree with the contents of this manuscript.

ADDITIONAL INFORMATION

REFERENCES

1. Teras, L. R., DeSantis, C. E., Cerhan, J. R., Morton, L. M., Jemal, A. & Flowers, C. R. 2016 US lymphoid malignancy statistics by World Health Organization subtypes. *CA Cancer J. Clin.* **66**, 443–459 (2016).
2. Gu, X., Zheng, R., Xia, C., Zeng, H., Zhang, S., Zou, X. et al. Interactions between life expectancy and the incidence and mortality rates of cancer in China: a population-based cluster analysis. *Cancer Commun.* https://doi.org/10.1186/s40880-018-0308-x (2018).
3. Pfreundschuh, M., Kuhnt, E., 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.* **12**, 1013–1022 (2011).
4. 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* **116**, 2040–2045 (2010).
5. Shipp, M. A., Harrington, D. P., Anderson, J. R., Armitage, J. O., Bonadonna, G., Brittinger, G. et al. A predictive model for aggressive non-Hodgkins-lymphoma. *N. Engl. J. Med.* **329**, 987–994 (1993).
6. Zhou, Z., Sehn, L. H., Rademaker, A. W., Gordon, L. I., LaCasce, A. S., Crosby-Thompson, A. et al. An enhanced International Prognostic Index (NCCN-IPI) for patients with diffuse large B-cell lymphoma treated in the rituximab era. *Blood* **123**, 837–842 (2014).
7. Sehn, L. H., Berry, B., Chhanabhai, M., Fitzgerald, C., Gill, K., Hoskins, P. 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**, 1857–1861 (2007).
8. Prochazka, K. T., 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* **115**, 1264–1272 (2016).
9. Montalban, C., Diaz-Lopez, A., Dlouhy, I., Rovira, J., Lopez-Guillermo, A., Alonso, S. et al. Validation of the NCCN-IPI for diffuse large B-cell lymphoma (DLBCL): the addition of beta(2)-microglobulin yields a more accurate GELTAMO-IPI. *Br. J. Haematol.* **176**, 918–928 (2017).
10. Schmitz, R., Wright, G. W., Huang, D. W., Johnson, C. A., Phelan, J. D., Wang, J. Q. et al. Genetics and pathogenesis of diffuse large B-cell lymphoma. *N. Engl. J. Med.* **378**, 1396–1407 (2018).
11. Chapuy, B., Stewart, C., Dunford, A. J., Kim, J., Kamburov, A., Redd, R. A. et al. Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. *Nat. Med.* **24**, 679–690 (2018).
12. Oh, S. Y., Kim, W. S., Kim, J. S., Kim, S. J., Yoon, D. H., Yang, D. H. et al. Phase II study of R-CVP followed by rituximab maintenance therapy for patients with advanced marginal zone lymphoma: consortium for improving survival of lymphoma (CISL) study. *Cancer Commun.* https://doi.org/10.1186/s40880-019-0403-7 (2019).
13. Wright, G. W., Huang, D. W., Phelan, J. D., Coulibaly, Z. A., Roulland, S., Young, R. M. et al. A probabilistic classification tool for genetic subtypes of diffuse large B cell lymphoma with therapeutic implications. *Cancer Cell* **37**, 551–568 (2020).

14. Yang, R., Shao, T., Long, M., Shi, Y., Liu, Q., Yang, L. et al. Long noncoding RNA PVT1 promotes tumor growth and predicts poor prognosis in patients with diffuse large B-cell lymphoma. *Cancer Commun.* **40**, 551–555 (2020).

15. Bari, A., Marcheselli, L., Sacchi, S., Marcheselli, R., Pozzi, S., Ferri, P. et al. Prognostic models for diffuse large B-cell lymphoma in the rituximab era: a never-ending story. *Ann. Oncol.* **21**, 1486–1491 (2010).

16. Chen, Y., Neelapu, S., Feng, L., Bi, W., Yang, T.-H., Wang, M. et al. Prognostic significance of baseline peripheral absolute neutrophil, monocyte and serum 2-microglobulin level in patients with diffuse large b-cell lymphoma: a new prognostic model. *Br. J. Haematol.* **175**, 290–299 (2016).

17. Kanemasa, Y., Shimoyama, T., Sasaki, Y., Tamura, M., Sawada, T., Omuro, Y. et al. Beta-2 microglobulin as a significant prognostic factor and a new risk model for patients with diffuse large B-cell lymphoma. *Hematol. Oncol.* **35**, 440–446 (2017).

18. Kim, D. H., Baek, J. H., Chae, Y. S., Kim, Y. K., Kim, H. J., Park, Y. H. et al. Absolute lymphocyte counts predicts response to chemotherapy and survival in diffuse large B-cell lymphoma. *Leukemia* **21**, 2227–2230 (2007).

19. Wilcox, R. A., Ristow, K., Habermann, T. M., Inwards, D. J., Micallef, I. N. M., Johnston, P. B. et al. The absolute monocyte and lymphocyte prognostic score predicts survival and identifies high-risk patients in diffuse large-B-cell lymphoma. *Leukemia* **25**, 1502–1509 (2011).

20. 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* **130**, 489–500 (2017).

21. Johnson, N. A., Slack, G. W., Savage, K. J., Connors, J. M., Ben-Neriah, S., Rogic, S. et al. Concurrent expression of MYC and BCL2 in diffuse large B-cell lymphoma treated with rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone. *J. Clin. Oncol.* **30**, 3452–3459 (2012).

22. Hu, S., Xu-Monette, Z. Y., Tzankov, A., Green, T., Wu, L., Balasubramanyam, A. et al. MYC/BCL2 protein coexpression contributes to the inferior survival of activated B-cell subtype of diffuse large B-cell lymphoma and demonstrates high-risk gene expression signatures: a report from The International DLBCL Rituximab-CHOP Consortium Program. *Blood* **121**, 4021–4031 (2013).

23. Yamaguchi, M., Seto, M., Okamoto, M., Ichinohasama, R., Nakamura, N., Yoshino, T. et al. De novo CD5(+) diffuse large B-cell lymphoma: a clinicopathologic study of 109 patients. *Blood* **99**, 815–821 (2002).

24. Niitsu, N., Okamoto, M., Tamaru, J. I., Yoshino, T., Nakamura, N., Nakamura, S. et al. Clinicopathologic characteristics and treatment outcome of the addition of rituximab to chemotherapy for CD5-positive in comparison with CD5-negative diffuse large B-cell lymphoma. *Ann. Oncol.* **21**, 2069–2074 (2010).

25. Scott, D. W., Mottok, A., Ennishi, D., Wright, G. W., Farinha, P., Ben-Neriah, S. et al. Prognostic significance of diffuse large B-cell lymphoma cell of origin determined by digital gene expression in formalin-fixed paraffin-embedded tissue biopsies. *J. Clin. Oncol.* **33**, 2848–2856 (2015).

26. Wight, J. C., Chong, G., Grigg, A. P. & Hawkes, E. A. Prognostication of diffuse large B-cell lymphoma in the molecular era: moving beyond the IPI. *Blood Rev.* **32**, 400–415 (2018).

27. Staiger, A. M., Ziepert, M., Horn, H., Scott, D. W., Barth, T. F. E., Bernd, H.-W. et al. Clinical impact of the cell-of-origin classification and the MYC/BCL2 dual expresser status in diffuse large B-cell lymphoma treated within prospective clinical trials of the German High-grade Non-Hodgkin's Lymphoma Study Group. *J. Clin. Oncol.* **35**, 2515–2526 (2017).

28. Yang, Y., Zhang, Y. J., Zhu, Y., Cao, J. Z., Yuan, Z. Y., Xu, L. M. et al. Prognostic nomogram for overall survival in previously untreated patients with extranodal NK/T-cell lymphoma, nasal-type: a multicenter study. *Leukemia* **29**, 1571–1577 (2015).

29. Tian, X. P., Huang, W. J., Huang, H. Q., Liu, Y. H., Wang, L., Zhang, X. et al. Prognostic and predictive value of a microRNA signature in adults with T-cell lymphoblastic lymphoma. *Leukemia* **33**, 2454–2465 (2019).

30. Liang, W., Zhang, L., Jiang, G., Wang, Q., Liu, L., Liu, D. et al. Development and validation of a nomogram for predicting survival in patients with resected non-small-cell lung cancer. *J. Clin. Oncol.* **33**, 861–869 (2015).

31. Swerdlow, S. H., Campo, E., Pileri, S. A., Harris, N. L., Stein, H., Siebert, R. et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood* **127**, 2375–2390 (2016).

32. Cheson, B. D., Fisher, R. I., Barrington, S. F., Cavalli, F., Schwartz, L. H., Zucca, E. et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. *J. Clin. Oncol.* **32**, 3059–3068 (2014).

33. Weiss, L., Melchardt, T., Habringer, S., Boekstegers, A., Hufnagl, C., Neureiter, D. et al. Increased body mass index is associated with improved overall survival in diffuse large B-cell lymphoma. *Ann. Oncol.* **25**, 171–176 (2014).

34. Cao, Y., Shi, Y.-X., Chen, J.-O., Tan, Y.-T., Cai, Y.-C., Luo, H.-Y. et al. Serum C-reactive protein as an important prognostic variable in patients with diffuse large B cell lymphoma. *Tumor Biol.* **33**, 1039–1044 (2012).

35. Ennishi, D., Takeuchi, K., Yokoyama, M., Asai, H., Mishima, Y., Terui, Y. et al. CD5 expression is potentially predictive of poor outcome among biomarkers in patients with diffuse large B-cell lymphoma receiving rituximab plus CHOP therapy. *Ann. Oncol.* **19**, 1921–1926 (2008).

36. Hans, C. P., Weisenburger, D. D., Greiner, T. C., Gascoyne, R. D., Delabie, J., Ott, G. et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood* **103**, 275–282 (2004).

37. Tadmor, T., Bari, A., Sacchi, S., Marcheselli, L., Liardo, E. V., Avivi, I. et al. Monocyte count at diagnosis is a prognostic parameter in diffuse large B-cell lymphoma: results from a large multicenter study involving 1191 patients in the pre- and post-rituximab era. *Haematologica* **99**, 125–130 (2014).

38. Lin, B., Chen, C., Qian, Y. & Feng, J. Prognostic role of peripheral blood lymphocyte/monocyte ratio at diagnosis in diffuse large B-cell lymphoma: a meta-analysis. *Leuke Lymphoma* **56**, 2563–2568 (2015).

39. Markovic, O., Popovic, L., Marisavljevic, D., Jovanovic, D., Filipovic, B., Stanisavljevic, D. et al. Comparison of prognostic impact of absolute lymphocyte count, absolute monocyte count, absolute lymphocyte count/absolute monocyte count prognostic score and ratio in patients with diffuse large B cell lymphoma. *Eur. J. Intern. Med.* **25**, 296–302 (2014).

40. Wei, X., Huang, F., Wei, Y., Jing, H., Xie, M., Hao, X. et al. Low lymphocyte-to-monocyte ratio predicts unfavorable prognosis in non-germinal center type diffuse large B-cell lymphoma. *Leuke. Res.* **38**, 694–698 (2014).

41. Morschhauser, F., Feugier, P., Flinn, I. W., Gasiorowski, R. E., Greil, R., Illes, A. et al. Venetoclax plus rituximab, cyclophosphamide, doxorubicin, vincristine and prednisolone (R-CHOP) improves outcomes in BCL2-positive first-line diffuse large B-cell lymphoma (DLBCL): first safety, efficacy and biomarker analyses from the phase II CAVALLI study. *Blood* **132**, 5 (2019). (abstract 782).

42. Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A. et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* **403**, 503–511 (2000).

43. Rosenwald, A., Wright, G., Chan, W. C., Connors, J. M., Campo, E., Fisher, R. I. et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N. Engl. J. Med.* **346**, 1937–1947 (2002).

44. Cai, Q.-Q., Hu, L.-Y., Geng, Q.-R., Chen, J., Lu, Z.-H., Rao, H.-L. et al. New risk factors and new tendency for central nervous system relapse in patients with diffuse large B-cell lymphoma: a retrospective study. *Chin. J. Cancer* https://doi.org/10.1186/s40880-016-0150-y (2016).

45. Cai, Q., Fang, Y. & Young, K. H. Primary central nervous system lymphoma: molecular pathogenesis and advances in treatment. *Transl. Oncol.* **12**, 523–538 (2019).

46. Schmitz, N., Zeynalova, S., Nickelsen, M., Kansara, R., Villa, D., Sehn, L. H. et al. CNS international prognostic index: a risk model for CNS relapse in patients with diffuse large B-cell lymphoma treated with R-CHOP. *J. Clin. Oncol.* **34**, 3150–3156 (2016).