Cancer Causes Control (2015) 26:1163–1172
DOI 10.1007/s10552-015-0610-8

ORIGINAL PAPER

Validation of clinical prognostic indices for diffuse large B-cell lymphoma in the National Cancer Data Base

Adam J. Olszewski[1,2] · **Eric S. Winer**[1,3] · **Jorge J. Castillo**[4,5]

Received: 24 January 2015 / Accepted: 2 June 2015 / Published online: 9 June 2015

Abstract
Background Accurate risk stratification is necessary for epidemiologic and outcomes research in diffuse large B-cell lymphoma (DLBCL). We evaluated performance characteristics of the clinically derived International Prognostic Index (IPI) and revised IPI (R-IPI) with a regression model-based score using the National Cancer Data Base.
Methods We studied DLBCL patients diagnosed in 2004–2011, divided into derivation and validation cohorts. The model-based score was calculated from a Cox model incorporating variables routinely recorded by cancer registries. Calibration and discrimination of the indices with regard to overall survival were evaluated in the validation cohort.
Results The IPI was recorded in 19,511 of 119,942 patients, of whom 79 % received chemotherapy. Both clinical indices provided good survival discrimination (5-year estimate range 33–74 % for the IPI, and 41–87 % for the R-IPI), but explained only 16 % of variation in survival. Survival predictions among chemotherapy-treated patients were similar to estimates from published clinical cohorts. The model-based score had significantly better discrimination characteristics (5-year survival estimate range 22–87 %) and explained 23 % of variation in survival.
Conclusions We validated the IPI and R-IPI as recorded by cancer registries to provide robust risk stratification in the general population with DLBCL, but a prognostic model using raw registry data provides superior performance. Explicit recording of prognostic factors is preferable to abstracting coarsened clinical indices for the purpose of population-based epidemiologic research. Considering low variation of survival explained by the standard clinical variables, incorporating molecular markers into registry data is necessary to improve risk stratification.

Keywords Diffuse large B-cell lymphoma · Non-Hodgkin lymphoma · Prognosis · Epidemiology · Outcomes research · Survival analysis · Model validation

Preliminary data from this study were presented at the 56th American Society of Hematology Annual Meeting, December 6–9, 2014, San Francisco, CA.

✉ Adam J. Olszewski
adam_olszewski@brown.edu

[1] Department of Medicine, Alpert Medical School of Brown University, Providence, RI, USA

[2] Division of Hematology-Oncology, Memorial Hospital of Rhode Island, 111 Brewster St., Pawtucket, RI 02860, USA

[3] Division of Hematology-Oncology, Rhode Island Hospital, Providence, RI, USA

[4] Department of Medicine, Harvard Medical School, Boston, MA, USA

[5] Division of Hematologic Oncology, Dana-Farber Cancer Institute, Boston, MA, USA

Introduction

Research using population-derived data can generate important knowledge about the effectiveness and toxicity of cancer therapy, heterogeneity of treatment effects, costs, and access to treatment, thus answering questions difficult to address in randomized clinical trials [1]. Because of treatment indication bias characteristic of observational studies, validated risk stratification is critically important when survival outcomes between subpopulations are compared [2]. For non-Hodgkin lymphoma (NHL), and specifically for its most common aggressive subtype diffuse large B-cell lymphoma (DLBCL), prognosis in clinical practice is determined using scores that integrate

patient- and disease-related factors. The classic International Prognostic Index (IPI), developed in 1993, identified age over 60 years, poor performance status, stage III/IV lymphoma, elevated serum lactate dehydrogenase (LDH), and involvement of two or more extranodal sites as high-risk factors. The IPI stratified DLBCL into four groups with 5-year overall survival (OS) ranging from 26 to 73 % [3]. Subsequent studies performed in the era of modern rituximab-based immunochemotherapy showed a better risk discrimination when the IPI was re-categorized into a three-level index (termed revised IPI, R-IPI) [4]. Other reports provided conflicting assessments of the IPI and R-IPI after treatment with rituximab-based immunochemotherapy [5–7]. Recently, an enhanced clinical predictor (the National Comprehensive Cancer Network-IPI, NCCN-IPI) was developed by subdividing age and LDH levels into finer categories, and designating specific high-risk extranodal sites [8].

All iterations of the IPI were derived from clinical trial cohorts uniformly treated with anthracycline-based chemotherapy. Until now, their performance in population-based data has not been evaluated. Patients enrolled in clinical trials have different characteristics than the DLBCL population at large [9]. In fact, up to 40 % of Medicare enrollees with DLBCL may not receive standard outpatient immunochemotherapy [10]. Patients aged >80 years or with comorbidities are at particularly high risk of not receiving curative treatment [11]. The objective of our study was to evaluate performance of the IPI and R-IPI with regard to survival prediction in a population-wide DLBCL cohort and to compare those clinically derived indices with a score based on raw variables routinely recorded by cancer registries. Our rationale was to identify the optimal risk stratification method for epidemiologic and outcomes research on lymphomas using population data and to assess whether abstracting clinical indices by cancer registries is preferable to raw recording of prognostic factors.

Materials and methods

Data source and cohort selection

This study used the de-identified National Cancer Data Base (NCDB) and was exempt from oversight by the institutional human subject protection committee. The NCDB is a joint project of the Commission on Cancer of the American College of Surgeons and the American Cancer Society. It collects registry records from over 1,500 cancer programs, capturing approximately 70 % of all incident cancer cases in the United States of America (USA) [12]. We selected DLBCL patients diagnosed between 2004 and 2011 using histology codes according to the InterLymph classification of lymphomas for epidemiologic research [13]. The NCDB files contained patient demographics, primary site and stage of the lymphoma, presence or absence of B symptoms, and the first course of treatment, including indicator of chemotherapy receipt. Specific chemotherapy regimens were not recorded. Socioeconomic status was approximated by median income in the patient's zip code according to the 2012 American Community Survey [14].

Since 2004, cancer registries in the USA have been directly recording the IPI as a lymphoma-specific factor in the Collaborative Stage Data Collection System. Values of the specific components (LDH, performance status) are not collected, only the summary score. The IPI scores must be explicitly documented by treating physicians in the medical record. We divided the cohort into two subgroups depending on whether the IPI was or was not recorded. Patients whose IPI was not available were used as the *derivation cohort* for the regression model-based score. Comparative evaluation of all prognostic indices was then conducted in the *validation cohort* of cases with recorded IPI. The IPI was classified into standard-risk groups designated as low (L, score 0 or 1), low intermediate (LI, 2), high intermediate (HI, 3) and high (H, 4 or 5) [3]. The R-IPI was stratified into "very good" (score 0), "good" (1 or 2), or "poor" (3–5) categories [4]. OS was calculated from the date of diagnosis. Disease-specific end points like event-free survival were not available because recurrences of lymphoma were not recorded.

Development of the model-based prognostic score

We developed a Cox survival model in the derivation cohort using available variables routinely recorded by the cancer registries and known to associate with survival in DLBCL, namely: age, sex, race, socioeconomic status, stage, extranodal primary site, B symptoms, and Charlson–Deyo comorbidity score, which maps diagnoses from patient's medical record to comorbidities predictive of increased mortality [15]. High-risk extranodal primary sites (central nervous system, lung, liver, pancreas, gastrointestinal tract, and bone marrow) were further distinguished based on prior epidemiologic studies [8, 16]. Because a nonlinear relationship between age and OS was evident, we transformed age using a fractional polynomial—a method that avoids model misspecification and information loss occurring when a continuous variable is coarsely categorized [17]. The proportional hazard assumption was assessed graphically using plots of scaled Schoenfeld residuals against time, but statistical testing for violations was not used because of the large sample size. We subsequently converted the Cox model into a prognostic score and validated it according to the principles delineated by Royston and Altman [18]. Cutoffs for four risk groups (termed L, LI, HI, H in analogy to the IPI) were established within the derivation cohort at the mean and ± 1 SD of the model linear predictor. These cutoffs correspond to

the 16th, 50th, and 84th percentiles of mortality risk. Using the model coefficients (β_1–β_N) listed in Table 2 and values of each variable (X_1–X_N), the linear predictor ($X\beta$) of the model can be then calculated for any patient as:

$$X\beta = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_N X_N$$

Using the same coefficients and cutoffs (i.e., without refitting the model), the model-based score was then calculated in the validation cohort. Concordance between the IPI and the model-based score was evaluated using Cohen's κ coefficient. This measure of inter-rater agreement accounts for chance agreement and varies between 0 (no agreement) and 1 (perfect agreement).

Evaluation of the prognostic indices

We compared the prognostic indices with regard to their *calibration*, defined as consistent accuracy of survival prediction, and *discrimination*, defined as degree of survival curve separation between risk groups, in the validation cohort, in which all cases had a recorded IPI score. For calibration, OS estimates for those groups were compared against published data. In order to adequately match the NCDB population with the clinical trial cohorts used for the IPI derivation, these comparisons were conducted for patients who received multi-agent chemotherapy and excluding cases with primary central nervous system lymphoma (PCNSL) or with recorded human immunodeficiency virus (HIV) infection. For the calibration assessment of the model-based score, OS was compared between our NCDB validation and derivation cohorts.

Discrimination was assessed graphically using Kaplan–Meier curves and by two statistics. The concordance probability estimate (CPE) measures the proportion of all patient pairs with concordant predictions and outcomes (meaning that the patient from a higher-risk group has a shorter observed survival time) [19]. Royston and Sauerbrei [18, 20] explained variation statistic (R_D^2) measures the amount of variation in survival explained by a prognostic index and reflects its overall predictive ability. We calculated 95 % confidence intervals (CI) for those statistics using the bootstrap method with 1,000 replications. Continuous variables were compared with the Wilcoxon rank-sum test and categorical variables with the Chi-squared test. All analyses were performed using Stata/SE 13.1 (StataCorp LP, College Station, TX, USA).

Results

Patient characteristics

We identified 119,942 patients with DLBCL, of whom 79 % received chemotherapy. The IPI was recorded by the registry in only 16.3 % of cases. Patients whose IPI was recorded had a lower rate of other missing data and a higher rate of chemotherapy utilization (Table 1). We found manifestly erroneous IPI values in 1,890 (9.7 %) of records (for example, when the IPI score was 0 while age was >60 years) and corrected them prior to analysis. Median follow-up of censored patients was 47.5 months (interquartile range 27–72 months), median OS in the entire cohort (Fig. 1a) was 5.7 years (CI 5.6–5.8), and 5-year OS estimate was 52.0 % (CI 51.6–52.3). Patients who received chemotherapy had a median OS of 7.5 years (CI 7.3–7.7) with 5-year OS of 58 % (CI 57.6–58.4), while untreated patients had a median OS of only 0.4 years (CI 0.3–0.4).

Comparison of the prognostic indices

The Cox regression model for OS (Table 2) used for development of the model-based score accounted for the complex relationship between hazard ratio and age (Fig. 1b). The cutoffs for the model linear predictor ($X\beta$) were established at ≤0.4276 (L), 0.4276–1.1343 (LI), 1.1343–1.9229 (HI), and >1.9229 (H). The model-based score only partially overlapped with the IPI, with absolute concordance of 39 % and Cohen's κ of 0.19, indicating poor inter-rater agreement.

Proportions of DLBCL cases falling into the IPI-defined L, LI, HI and H groups in the NCDB population (Table 3) were similar to the original International NHL Prognostic Factors Project study (35, 27, 22, and 16 %, respectively) [3]. The percentage of cases designated by R-IPI as "very good," "good," or "poor" also matched the original British Columbia Cancer Agency (BCCA) study (10, 45 and 45 %, respectively) [4]. The model-based score, according to its design, allocated cases to four risk groups at predefined proportions (16, 34, 34 and 16 %) in the derivation cohort, which were then consistent in the validation cohort.

Calibration of the survival prediction

OS in the NCDB validation cohort was significantly lower within all IPI- or R-IPI-defined strata compared with published values from training datasets. However, estimates calculated for the subpopulation of patients ($n = 14{,}986$) who received multi-agent chemotherapy (and excluding HIV-associated DLBCL and PCSNL) were closer: 5-year OS in the NCDB was 80, 63, 56, and 41 % for the IPI L, LI, HI, and H risk groups, respectively. In comparison, analogous estimates had been reported as 90, 77, 62, and 54 % in the NCCN dataset, and 84, 72, 54, and 43 % in the BCCA dataset, respectively [8]. Compared with the R-IPI derivation study (4-year OS of 94, 79 and 55 % for "very good," "good," and "poor" risk groups, respectively), the NCDB estimates were also close at 92, 74 and 52 %, respectively,

Table 1 Characteristics of patients with diffuse large B-cell lymphoma from the National Cancer Data Base

Variable	Training cohort		Validation cohort		*p*	All cases	
	n	%	*n*	%		*n*	%
Total number of cases	100,431		19,511			119,942	
Age, median (interquartile range)	68 (55–78)		67 (55–77)		<0.001	67 (55–78)	
Age groups							
18–39 years	7,470	7.4	1,459	7.5	<0.001	8,929	7.4
40–59 years	25,650	25.5	5,114	26.2		30,764	25.6
60–74 years	33,642	33.5	6,723	34.5		40,365	33.7
≥75 years	33,669	33.5	6,215	31.9		39,884	33.3
Sex							
Women	46,699	46.5	9,158	46.9	0.26	55,857	46.6
Men	53,732	53.5	10,353	53.1		64,085	53.4
Race/ethnicity							
White non-Hispanic	80,720	80.4	16,641	85.3	<0.001	97,361	81.2
White Hispanic	6,662	6.6	879	4.5		7,541	6.3
Black	8,031	8.0	1,188	6.1		9,219	7.7
American Indian	222	0.2	60	0.3		282	0.2
Asian	2,848	2.8	475	2.4		3,323	2.8
Unrecorded	1,948	1.9	268	1.4		2,216	1.8
Ann Arbor stage							
I	28,237	28.1	3,928	20.1	<0.001	32,165	26.8
II	18,593	18.5	3,845	19.7		22,438	18.7
III	16,022	16.0	3,934	20.2		19,956	16.6
IV	33,536	33.4	7,695	39.4		41,231	34.4
Unrecorded	4,043	4.0	109	0.6		4,152	3.5
B symptoms							
Absent	58,927	58.7	12,004	61.5	<0.001	70,931	59.1
Present	25,995	25.9	6,108	31.3		32,103	26.8
Unrecorded	15,509	15.4	1,399	7.2		16,908	14.1
Primary site							
Nodal	64,463	64.2	13,581	69.6	<0.001	78,044	65.1
High-risk extranodal[b]	16,879	16.8	2,992	15.3		19,871	16.6
Other extranodal	19,089	19.0	2,938	15.1		22,027	18.3
Charlson–Deyo Score							
0	75,081	74.8	14,193	72.7	<0.001	89,274	74.4
1	16,732	16.7	3,666	18.8		20,398	17.0
≥2	8,618	8.6	1,652	8.5		10,270	8.6
Median income, by zip code							
<$38,000	16,609	16.5	3,000	15.4	<0.001	19,609	16.3
$38,000–$47,999	23,217	23.1	4,971	25.5		28,188	23.5
$48,000–$62,999	26,630	26.5	5,343	27.4		31,973	26.7
≥$63,000	31,320	31.2	5,741	29.4		37,061	30.9
Unknown zip code	2,655	2.6	456	2.3		3,111	2.6
Facility type							
Community program[a]	10,592	10.5	2,397	12.3	<0.001	12,989	10.8
Comprehensive community program	53,485	53.3	10,695	54.8		64,180	53.5
Academic/research program	36,354	36.2	6,419	32.9		42,773	35.7
Chemotherapy							
Administered	77,822	77.5	16,918	86.7	<0.001	94,740	79.0

Table 1 continued

Variable	Training cohort		Validation cohort		*p*	All cases	
	n	%	*n*	%		*n*	%
Not administered	20,554	20.5	2,423	12.4		22,977	19.2
Unknown	2,055	2.0	170	0.9		2,225	1.9

[a] Including 119 cases from other types of programs

[b] Central nervous system, lung, liver, pancreas, gastrointestinal tract, and bone marrow

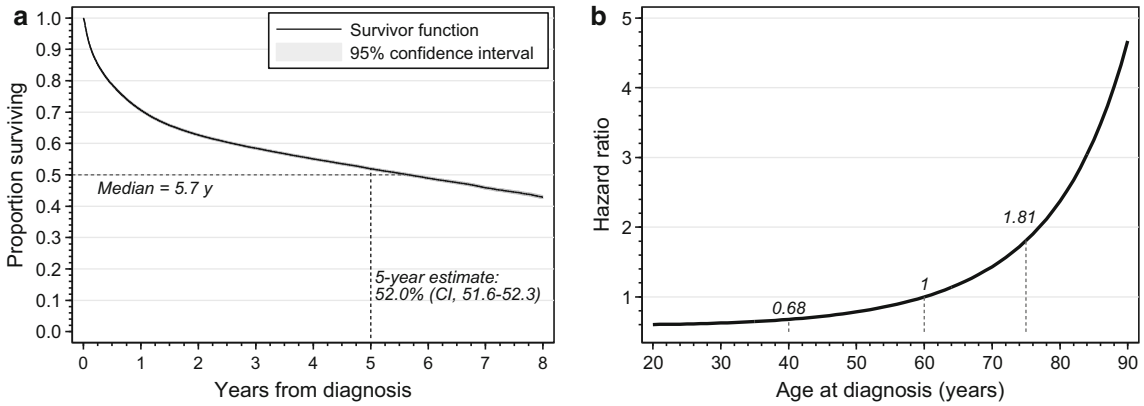

Fig. 1 **a** Overall survival curve for all patients with diffuse large B-cell lymphoma in the National Cancer Data Base (2004–2011). **b** The nonlinear relationship between age and relative hazard of death in patients with diffuse large B-cell lymphoma. The relationship is represented by a fractional polynomial *curve* centered at reference age of 60 years; values of relative hazard at age 40 and 75 are additionally shown

for patients treated with chemotherapy [4]. For the model-based score, 5-year OS in the entire validation cohort (as listed in Table 3) was 2–5 % higher than in the derivation cohort (84, 63, 39, and 18 % for L, LI, HI, and H risk groups, respectively).

Survival discrimination

The statistical measures of survival discrimination showed similar values for the IPI and R-IPI and were significantly better for the model-based score in every aspect: graphical separation of the survival curves (Fig. 2), hazard ratios for risk groups, and both CPE and R_D^2. The explained variation of survival (R_D^2) was only 16 % for the IPI and R-IPI, although even the regression model-based score explained only 26 % of this variation. When our main Cox model was further augmented by the recorded IPI score (thus incorporating information on LDH and patient's performance status), the resulting model achieved R_D^2 of 29.6 %. The comparisons of indices yielded similar results in the subcohort of patients who received multi-agent chemotherapy, with CPE of 65.5, 62.8, and 67.3 % for the IPI, R-IPI and model-based score, respectively, and R_D^2 of 16.8, 16.8, and 22.6 %, respectively.

Discussion

To our knowledge, this study is the first evaluation of commonly used clinical prognostic indices in a data source from an unselected US population, although they were studied in several large datasets [4–6, 8, 21, 22]. We applied the IPI and R-IPI in a comprehensive cohort containing a majority of DLBCL cases diagnosed in the USA between 2004 and 2011, regardless of treatment status. Although survival in this unselected cohort was naturally lower than in the reference published datasets, both scores were validated to provide robust risk discrimination. Not surprisingly, they were better calibrated in the subcohort of patients treated with chemotherapy. Nevertheless, a data-driven, model-based score showed significantly better characteristics. Therefore, for maximum efficiency in risk stratification and bias adjustment, research projects using population-based data should employ properly specified multivariable regression models on raw recorded variables rather than attempt to replicate the IPI groupings.

The IPI was designed using data from patients with various aggressive lymphomas (B cell and T cell) treated in the 1980s [3]. Studies assessing its utility for DLBCL in the era of modern immunochemotherapy reported conflicting results, although it remains widely used for stratification in prospective clinical trials and for daily clinical practice.

Table 2 Cox model for overall survival in the derivation cohort of 100,431 patients with diffuse large B-cell lymphoma

Variable	Coefficient[a] (β)	Hazard ratio (e^{β})	95 % confidence interval	p
Age[b]				
Term 1: $(age/10)^3$	−0.001182			<0.00001
Term 2: $(age/10)^3 \times \log(age/10)$	0.001841			
Sex				
Female		Reference		
Male	0.140678	1.15	1.13–1.17	<0.00001
Race				
White non-Hispanic		Reference		
White Hispanic	0.029176	1.03	0.99–1.07	0.151
Black	0.226866	1.25	1.21–1.30	<0.00001
American Indian	0.124525	1.13	0.93–1.38	0.21
Asian	0.064064	1.07	1.01–1.13	0.032
Unrecorded	−0.046583	0.95	0.89–1.03	0.22
Ann Arbor stage				
I		Reference		
II	−0.035120	0.97	0.94–1.00	0.025
III	0.171432	1.19	1.15–1.23	<0.00001
IV	0.493902	1.64	1.60–1.68	<0.00001
Unrecorded	0.436381	1.55	1.47–1.62	<0.00001
Primary site				
Nodal		Reference		
High-risk extranodal[c]	0.258710	1.30	1.27–1.33	<0.00001
Other extranodal	−0.250136	0.78	0.76–0.80	<0.00001
B symptoms				
Absent		Reference		
Present	0.226708	1.25	1.23–1.28	<0.00001
Unrecorded	0.141062	1.15	1.12–1.18	<0.00001
Charlson–Deyo Score				
0		Reference		
1	0.308782	1.36	1.33–1.39	<0.00001
2	0.741208	2.10	2.04–2.16	<0.00001
Median income				
<$38,000		Reference		
$38,000–$47,999	−0.089685	0.91	0.89–0.94	<0.00001
$48,000–$62,999	−0.138650	0.87	0.85–0.90	<0.00001
≥$63,000	−0.280196	0.76	0.73–0.78	<0.00001
Unrecorded	0.359973	1.43	1.36–1.51	<0.00001

[a] The linear predictor ($X\beta$) cutoffs for the model-based score risk groups were as follows: low: ≤0.4276; low intermediate: >0.4276 to ≤1.1343; high intermediate: >1.1343 to ≤1.9229; high: >1.9229

[b] Age was transformed as a fractional polynomial due to nonlinear association with survival. Only regression coefficients and summary P value for the fractional polynomial terms are listed

[c] Central nervous system, lung, liver, pancreas, gastrointestinal tract and bone marrow

Sehn et al. [4] found that the standard IPI applied to patients treated by current regimens identified only two discernible strata: L/LI (4-year OS 81–82 %) and HI/H (4-year OS 49–59 %), and proposed the R-IPI as a replacement. However, the "very good" R-IPI group contained only 10 % of patients, while the "poor" risk group contained 45 % of cases with a 4-year OS of 55 %, thus failing to identify the truly "poor" risk population. Similar characteristics of the R-IPI were reported in an Italian retrospective study, which further proposed absolute lymphocyte count as an additional

Table 3 Measures of risk stratification by prognostic indices in the validation cohort of 19,511 patients with diffuse large B-cell lymphoma

Index	Number of risk factors	% of cases	Survival at 5 years		Univariate		CPE (CI)	R_D^2 (CI)
			%	CI	Hazard ratio	CI		
IPI								
Low	0, 1	31.1	74.3	73.1–75.5	Reference			
Low intermediate	2	25.6	57.7	56.2–59.2	1.82	1.70–1.94	65.2 % (64.6–65.8)	16.3 % (15.2–17.5)
High intermediate	3	22.2	49.2	47.5–50.8	2.41	2.25–2.57		
High	4, 5	21.1	32.8	31.2–34.4	4.22	3.96–4.49		
Revised IPI								
Very good	0	7.2	87.2	85.1–89.0	Reference			
Good	1, 2	49.4	63.8	62.7–64.8	3.32	2.84–3.88	62.6 % (62.1–63.2)	16.2 % (15.0–17.5)
Poor	3–5	43.4	41.2	40.1–42.4	6.98	5.99–8.15		
Model-based score								
Low	[a]	16.0	86.9	85.5–88.1	Reference			
Low intermediate		34.4	67.9	66.6–69.1	2.82	2.54–3.15	69.0 % (68.4–69.5)	26.2 % (24.8–27.5)
High intermediate		33.8	44.8	43.4–46.1	5.86	5.28–6.51		
High		15.8	22.4	20.8–24.1	11.92	10.71–13.28		

CI confidence interval, *CPE* concordance probability estimate, *IPI* International Prognostic Index, R_D^2 explained variation statistic

[a] The model-based score was categorized according to the calculated linear predictor ($X\beta$) from the Cox model, using coefficients and cutoff values from Table 2

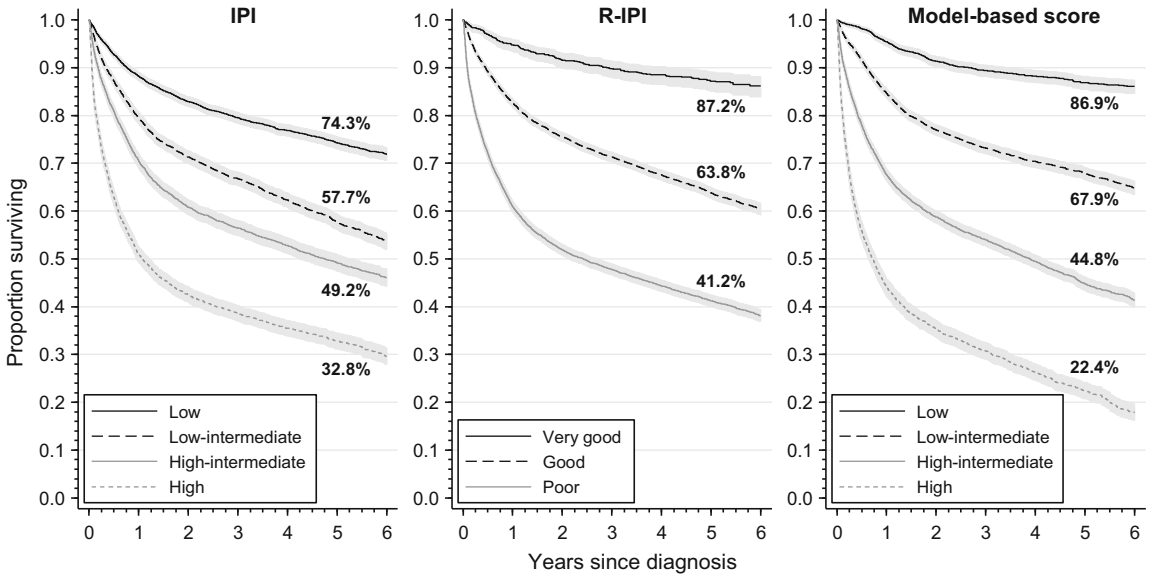

Fig. 2 Overall survival curves in the validation cohort, according to the risk group defined by the International Prognostic Index (IPI), revised IPI (R-IPI) or the model-based score. *Percentage values* indicate survival at 5 years; *bands* indicate 95 % confidence intervals

risk predictor [5]. In contrast, a combined analysis of three randomized immunochemotherapy trials (MabThera International Trial, MegaCHOEP, and RICOVER-60) indicated excellent performance characteristics of the classic IPI [6]. The enhanced NCCN-IPI index improved on the IPI specification by dividing age into four categories (≤40, 40–60, 60–75, and >75 years), LDH elevation into two levels (1–3 times, and >3 times upper limit of normal), and by identifying high-risk extranodal sites of involvement [8]. This new clinically useful index has been validated in a separate British Columbia Cancer Agency population-based database as part of the original study, showing higher CPE than the IPI (77 vs. 74 %), and confirmed its superiority in other subsequent DLBCL series (for example, CPE was 75 vs. 70 % in the study by Mian et al. [21–24]). Specific anatomical sites of lymphoma origin also have a prognostic value and may derive heterogeneous benefits from immunochemotherapy [10, 16].

In all iterations of the IPI, age >60 years was the strongest predictor of survival [3, 6]. To a large extent, this reflects the natural difference in mortality between younger and older patients. For example, in 2005 the death rate was 432 (per 100,000 person-years) for Americans aged 45–54 years and 2,137 for those aged 65–74 years [25]. The crude dichotomization of age at 60 years is a significant culprit of the IPI, particularly considering the median age at DLBCL diagnosis in the USA, which is 66 years [16]. Our analysis shows a nonlinear, exponential relationship between age and relative hazard of death in DLBCL. In large population-based data sets, when the goal is to adequately capture differences in delivery of chemotherapy and survival outcomes, this relationship must be modeled explicitly, without substantial information loss associated with age categorization [26]. Effective statistical methods to achieve this include fractional polynomials (as used in our study) or restricted cubic splines [17]. They require graphical presentation of the relationship because their associated regression coefficients have no direct interpretation. Descriptive summary of data in coarse age groups may thus still be of interest to clinicians. Additionally, our study corroborates recent findings from the California Cancer Registry showing that neighborhood socioeconomic status is associated with survival in DLBCL independently of purely clinical factors [27].

Remarkably, even when we combined the IPI and all available variables from the registry in a single, optimally specified model, the amount of explained survival variation in DLBCL remained below 30 %. Baseline clinical and sociodemographic factors may thus not be the main determinants of survival in this disease. Molecular markers of biological aggressiveness may be necessary for more precise risk stratification and prediction. Except for serum LDH, no such biomarkers are included in the IPI or the NCCN-IPI, although many have been discovered over the past decade. Identification of two gene expression clusters defining the germinal center B cell and activated B-cell DLBCL has been shown to improve prognostication within every IPI risk group [28, 29]. Accumulating evidence indicates that those two "cell-of-origin" subsets are driven by different oncogenic pathways and may respond to specific chemotherapeutic strategies, although the distinction has not yet been translated into widespread clinical practice [30, 31]. Recently, attention has been focused on the subset of DLBCL overexpressing the *MYC* gene in conjunction with *BCL2* and/or *BCL6* (referred to as double-hit or triple-hit lymphomas) [32]. This highly aggressive entity may have a particularly poor prognosis despite standard therapy [33, 34]. Many other immunophenotypical, genetic, and laboratory biomarkers are actively investigated for further risk stratification, as recently reviewed elsewhere [35, 36].

Our study has several limitations. We had to rely on the IPI values as abstracted by cancer registries and found evidence of erroneous IPI assignment in at least 10 % of cases. Other errors were undetected because of lack of direct assessment of LDH or performance status. We also could not unequivocally identify patients who received anthracycline-containing immunochemotherapy, because specific drugs and regimens were not recorded in the database. We were not able to evaluate to what extent any additional biomarkers could improve survival prediction in DLBCL, or whether physician-level or facility-level characteristics influenced patients' prognosis. Although our model-based score can be calculated for any patient with DLBCL using data from Table 2 and can produce a more precise survival prediction than the IPI/R-IPI, it is too complex for routine clinical practice, where weighted summary indices like the NCCN-IPI perform sufficiently well. We should emphasize that the model was developed to illustrate the best risk adjustment practices for population-based epidemiology and health outcomes research, and not for bedside use. The strengths of our study include the comprehensive nature of the cohort, which included adult patients of any age, and reliance on clinical registry records rather than administrative billing claims (as in e.g., studies utilizing the linked Surveillance, Epidemiology and End Results-Medicare data).

Our analysis suggests that in order to advance population-based epidemiologic and outcomes research, the effort of incorporating raw biomarkers into cancer registry data might be more of more value than abstracting coarsened scores like the IPI. Direct recording of serum LDH, some measure of performance status and involvement of the high-risk extranodal sites (which is already collected for solid tumors) might align the clinical content of the NCDB with the new NCCN-IPI clinical standard. Enhancements of the national registry programs with molecular prognostic markers have been advocated as necessary for robust research on comparative effectiveness of treatments or quality of health care and recently have been successfully piloted for breast cancer, colorectal cancer, and chronic myeloid leukemia [37–39]. The newly developed Hematopoietic and Lymphoid Database is an excellent platform to improve the usability of cancer registry data as a valuable research resource for studying hematologic malignancies if it can be augmented by validated biomarkers [40].

Acknowledgments The data used in the study are derived from a de-identified NCDB file. The American College of Surgeons and the Commission on Cancer have not verified and are not responsible for the analytic or statistical methodology employed, or the conclusions drawn from these data by the investigator. The study was funded by the authors' academic departments.

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

Ethical standard All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. For this type of study, formal consent is not required.

References

1. Meyer AM, Carpenter WR, Abernethy AP, Sturmer T, Kosorok MR (2012) Data for cancer comparative effectiveness research: past, present, and future potential. Cancer 118:5186–5197. doi:10.1002/cncr.27552
2. Korn EL, Freidlin B (2012) Methodology for comparative effectiveness research: potential and limitations. J Clin Oncol 30:4185–4187. doi:10.1200/JCO.2012.44.8233
3. The International Non-Hodgkin's Lymphoma Prognostic Factors Project (1993) A predictive model for aggressive non-Hodgkin's lymphoma. N Engl J Med 329:987–994. doi:10.1056/NEJM199309303291402
4. Sehn LH, Berry B, Chhanabhai M et al (2007) 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. doi:10.1182/blood-2006-08-038257
5. Bari A, Marcheselli L, Sacchi S et al (2010) Prognostic models for diffuse large B-cell lymphoma in the rituximab era: a never-ending story. Ann Oncol 21:1486–1491. doi:10.1093/annonc/mdp531
6. Ziepert M, Hasenclever D, Kuhnt E et al (2010) Standard International prognostic index remains a valid predictor of outcome for patients with aggressive CD20+ B-cell lymphoma in the rituximab era. J Clin Oncol 28:2373–2380. doi:10.1200/JCO.2009.26.2493
7. Advani RH, Chen H, Habermann TM et al (2010) Comparison of conventional prognostic indices in patients older than 60 years with diffuse large B-cell lymphoma treated with R-CHOP in the US Intergroup Study (ECOG 4494, CALGB 9793): consideration of age greater than 70 years in an elderly prognostic index (E-IPI). Br J Haematol 151:143–151. doi:10.1111/j.1365-2141.2010.08331.x
8. Zhou Z, Sehn LH, Rademaker AW et al (2014) An enhanced International Prognostic Index (NCCN-IPI) for patients with diffuse large B-cell lymphoma treated in the rituximab era. Blood 123:837–842. doi:10.1182/blood-2013-09-524108
9. Terschuren C, Gierer S, Brillant C, Paulus U, Loffler M, Hoffmann W (2010) Are patients with Hodgkin lymphoma and high-grade non-Hodgkin lymphoma in clinical therapy optimization protocols representative of these groups of patients in Germany? Ann Oncol 21:2045–2051. doi:10.1093/annonc/mdq214
10. Olszewski AJ, Winer ES, Castillo JJ (2014) Improved survival with rituximab-based chemoimmunotherapy in older patients with extranodal diffuse large B-cell lymphoma. Leuk Res 38:866–873. doi:10.1016/j.leukres.2014.04.009
11. Hamlin PA, Satram-Hoang S, Reyes C, Hoang KQ, Guduru SR, Skettino S (2014) Treatment patterns and comparative effectiveness in elderly diffuse large B-cell lymphoma patients: a surveillance, epidemiology, and end results-medicare analysis. Oncologist 19:1249–1257. doi:10.1634/theoncologist.2014-0113
12. Bilimoria KY, Stewart AK, Winchester DP, Ko CY (2008) The National Cancer Data Base: a powerful initiative to improve cancer care in the United States. Ann Surg Oncol 15:683–690. doi:10.1245/s10434-007-9747-3
13. Turner JJ, Morton LM, Linet MS et al (2010) InterLymph hierarchical classification of lymphoid neoplasms for epidemiologic research based on the WHO classification (2008): update and future directions. Blood 116:e90–e98. doi:10.1182/blood-2010-06-289561
14. American Community Survey (2012) United States Census Bureau. https://www.census.gov/acs/www/. Accessed 24 Jan 2015
15. Deyo RA, Cherkin DC, Ciol MA (1992) Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. J Clin Epidemiol 45:613–619. doi:10.1016/0895-4356(92)90133-8
16. Castillo JJ, Winer ES, Olszewski AJ (2014) Sites of extranodal involvement are prognostic in patients with diffuse large B-cell lymphoma in the rituximab era: an analysis of the Surveillance, Epidemiology and End Results database. Am J Hematol 89:310–314. doi:10.1002/ajh.23638
17. Binder H, Sauerbrei W, Royston P (2013) Comparison between splines and fractional polynomials for multivariable model building with continuous covariates: a simulation study with continuous response. Stat Med 32:2262–2277. doi:10.1002/sim.5639
18. Royston P, Altman DG (2013) External validation of a Cox prognostic model: principles and methods. BMC Med Res Methodol 13:33. doi:10.1186/1471-2288-13-33
19. Newson RB (2010) Comparing the predictive powers of survival models using Harrell's C or Somers' D. Stata J 10:339–358
20. Royston P (2006) Explained variation for survival models. Stata J 6:83–96
21. Mian M, Marcheselli L, Rossi A et al (2014) A diachronic-comparative analysis for the identification of the most powerful prognostic index for localized diffuse large B-cell lymphoma. Ann Oncol. doi:10.1093/annonc/mdu462
22. Troppan KT, Melchardt T, Deutsch A et al (2015) The significance of pretreatment anemia in the era of R-IPI and NCCN-IPI prognostic risk assessment tools: a dual-center study in diffuse large B-cell lymphoma patients. Eur J Haematol. doi:10.1111/ejh.12529
23. Huang CE, Chen YY, Lu CH, Chen PT, Lee KD, Chen CC (2015) Validation of an enhanced International Prognostic Index (NCCN-IPI) in an Asian cohort of patients with diffuse large B cell lymphoma. Ann Hematol. doi:10.1007/s00277-014-2293-8
24. Melchardt T, Troppan K, Weiss L et al (2015) A modified scoring of the NCCN-IPI is more accurate in the elderly and is improved by albumin and beta2-microglobulin. Br J Haematol 168:239–245. doi:10.1111/bjh.13116
25. Kung HC, Hoyert DL, Xu J, Murphy SL (2008) Deaths: final data for 2005. National vital statistics reports. National Center for Health Statistics, Hyattsville, MD
26. Simon R, Altman DG (1994) Statistical aspects of prognostic factor studies in oncology. Br J Cancer 69:979–985
27. Tao L, Foran JM, Clarke CA, Gomez SL, Keegan TH (2014) Socioeconomic disparities in mortality after diffuse large B-cell lymphoma in the modern treatment era. Blood 123:3553–3562. doi:10.1182/blood-2013-07-517110
28. Alizadeh AA, Eisen MB, Davis RE et al (2000) Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 403:503–511. doi:10.1038/35000501
29. Rosenwald A, Wright G, Chan WC et al (2002) The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. N Engl J Med 346:1937–1947. doi:10.1056/NEJMoa012914
30. Dunleavy K, Pittaluga S, Czuczman MS et al (2009) Differential efficacy of bortezomib plus chemotherapy within molecular subtypes of diffuse large B-cell lymphoma. Blood 113:6069–6076. doi:10.1182/blood-2009-01-199679
31. Sehn LH, Gascoyne RD (2015) Diffuse large B-cell lymphoma: optimizing outcome in the context of clinical and biologic heterogeneity. Blood 125:22–32. doi:10.1182/blood-2014-05-577189

32. Aukema SM, Siebert R, Schuuring E et al (2011) Double-hit B-cell lymphomas. Blood 117:2319–2331. doi:10.1182/blood-2010-09-297879
33. Petrich AM, Nabhan C, Smith SM (2014) MYC-associated and double-hit lymphomas: a review of pathobiology, prognosis, and therapeutic approaches. Cancer 120:3884–3895. doi:10.1002/cncr.28899
34. Dunleavy K (2014) Double-hit lymphomas: current paradigms and novel treatment approaches. ASH Educ Program Book 2014:107–112. doi:10.1182/asheducation-2014.1.107
35. Vaidya R, Witzig TE (2014) Prognostic factors for diffuse large B-cell lymphoma in the R(X)CHOP era. Ann Oncol 25:2124–2133. doi:10.1093/annonc/mdu109
36. Sehn LH (2012) Paramount prognostic factors that guide therapeutic strategies in diffuse large B-cell lymphoma. Hematol Am Soc Hematol Educ Program 2012:402–409. doi:10.1182/asheducation-2012.1.402
37. Altekruse SF, Rosenfeld GE, Carrick DM et al (2014) SEER cancer registry biospecimen research: yesterday and tomorrow. Cancer Epidemiol Biomark Prev 23:2681–2687. doi:10.1158/1055-9965.EPI-14-0490
38. Chen VW, Eheman CR, Johnson CJ et al (2014) Enhancing cancer registry data for comparative effectiveness research (CER) project: overview and methodology. J Regist Manag 41:103–112
39. Howlader N, Altekruse SF, Li CI et al (2014) US incidence of breast cancer subtypes defined by joint hormone receptor and HER2 status. J Natl Cancer Inst. doi:10.1093/jnci/dju055
40. Ruhl J, Adamo M, Dickie L, Sun L, Johnson CH (2014) Hematopoietic and lymphoid neoplasm coding manual. National Cancer Institute, Bethesda, MD 20850-9765