RESEARCH LETTER

blood advances

TO THE EDITOR:

Real-world applicability of the International Metabolic Prognostic Index in DLBCL: a validation cohort study

Daisuke Ikeda,[1] Mitsuaki Oura,[1] Atsushi Uehara,[1] Rikako Tabata,[1] Kentaro Narita,[1] Masami Takeuchi,[1] Youichi Machida,[2] and Kosei Matsue[1]

[1]Division of Hematology/Oncology, Department of Medicine and [2]Department of Radiology, Kameda Medical Center, Chiba, Japan

The International Metabolic Prognostic Index (IMPI) is a recently proposed model that shows improved outcome prediction for patients with diffuse large B-cell lymphoma (DLBCL) compared with that of the International Prognostic Index (IPI).[1] The IMPI incorporates the total metabolic tumor volume (TMTV) with age and stage as continuous variables rather than in a dichotomized manner, which is less influenced by data-driven optimal cutoffs, thereby increasing generalizability.[2,3] However, the applicability of IMPI to real-world cohorts remains uncertain due to its model construction using hybrid data sets that largely depend on clinical trials. Furthermore, a model performance comparison of IMPI and the National Comprehensive Cancer Network IPI (NCCN-IPI) remains unreported.[4]

To validate the prognostic utility of the IMPI, we retrospectively analyzed patients with newly diagnosed DLBCL who were initially treated with rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP) or R-CHOP–like chemotherapy at Kameda Medical Center between 2007 and 2020. The exclusion criteria included the absence of pretreatment fluorine-18 fluorodeoxyglucose (^{18}FDG)–positron emission tomography/computed tomography or lack of FDG-avid lesions or treatment with a less intensive regimen than R-CHOP (supplemental Figure 1). We followed the methodology from the original study, setting the threshold for delineating lymphoma lesions for TMTV calculation at an absolute standardized uptake value (SUV) of ≥4.0 and calculating IMPI using the regression formula: 0.003077330 × (MTV below the median of 307.9 mL) − 0.002761985 × (MTV above the median of 307.9 mL) + 0.008092449 × age − 0.114645415 × stage 2 + 0.281141117 × stage 3 + 0.322247142 × stage 4.[1] Patients were grouped into 4 categories based on IMPI scores, which corresponded to the respective category sizes used in the IPI and NCCN-IPI systems. Given the expected difference in median TMTV values between this and the original studies, we recalculated the IMPI by replacing the values of 307.9 with our median ($IMPI_{median}$). In addition, a 3-group separation based on continuous IMPI scores was used, which divided the cohort into the lowest 60% (low), the middle 30% (intermediate), and the highest 10% (high) risk groups. We compared 3-year time to progression (TTP), progression-free survival (PFS), and overall survival (OS) among IMPI, IPI, and NCCN-IPI categories, assessing their predictive performance with Akaike information criterion (AIC) and Harrell concordance index (c-index). For detailed ^{18}FDG-positron emission tomography/computed tomography imaging procedures, TMTV quantification, and statistical analyses, refer to the supplemental Methods. This study was conducted in accordance with the Declaration of Helsinki and approved by our institutional review board (approval number: 22-095).

In total, 385 patients were analyzed. Table 1 shows the comparison of baseline characteristics between the 2 studies. Our cohort had a median age that was 10 years higher (72 [interquartile range, 64-79] vs 62 [interquartile range, 51-70]) and a higher frequency of impaired performance status (23.9% vs 11.6%; $P < .001$). Disease stages were nearly evenly distributed in this study, whereas two-thirds of

Submitted 13 November 2023; accepted 2 February 2024; prepublished online on *Blood Advances* First Edition 15 February 2024; final version published online 11 April 2024. https://doi.org/10.1182/bloodadvances.2023012165.

Presented in abstract form at the 65th annual meeting of the American Society of Hematology, San Diego, CA, 11 December 2023.

The data sets generated in this study are available upon request from the corresponding author, Daisuke Ikeda (dskikd.2409@gmail.com).

The full-text version of this article contains a data supplement.

Table 1. Comparison of baseline characteristics between this and the original study

	Current cohort (n = 385)	Original study (n = 1241)	*P* value
Age, median (IQR), y	72 (64-79)	62 (51-70)	NA
≤60, n (%)	69 (17.9)	557 (44.9)	<.001
>60, n (%)	316 (82.1)	684 (55.1)	-
>75, n (%)	137 (35.5)	Not presented	NA
Stage, n (%)			
Limited stage	170 (44.2)	432 (34.8)	.001
Advanced stage	215 (55.8)	809 (65.2)	-
LDH, n (%)			
<UNL	121 (31.4)	507 (40.9)	<.001
>UNL	264 (68.6)	734 (59.1)	-
WHO-PS, n (%)			
≤1	293 (76.1)	1097 (88.4)	<.001
>1	92 (23.9)	144 (11.6)	-
Number of EN lesions, n (%)			
≤1	286 (74.3)	817 (65.8)	.001
>1	99 (25.7)	424 (34.2)	-
TMTV, mL (median, IQR)	198.2 (54.5-547.2)	307.9 (77.6-838.9)	NA
IPI category, n (%)			
Low	95 (24.7)	402 (32.4)	.004
Low-intermediate	94 (24.4)	276 (22.2)	.408
High-intermediate	92 (23.9)	321 (25.9)	.461
High	104 (27.0)	242 (19.5)	.002
Outcomes, % (95% CI)			
3-y TTP	69.3 (64.1-74.0)	79.7 (72.1-77.0)	NA
3-y PFS	60.3 (55.0-65.1)	74.5 (72.1-77.0)	NA
3-y OS	72.0 (67.1-76.4)	81.8 (79.7-84.0)	NA

EN, extranodal; IQR, interquartile range; LDH, lactate dehydrogenase; NA, not assessed; UNL, upper normal limit; WHO-PS, World Health Organization-performance status.

patients in the original study had advanced disease ($P = .001$). This difference may have contributed to the TMTV level in our study being approximately two-thirds of that in the original study (median, 198.2 mL vs 307.9 mL). The receiver operating characteristic curve determined optimal predictive values for our cohort as ~230 mL for all 3 end points (supplemental Figure 2), which was closer to the median value of our cohort. There was a significant difference in the distribution of IPI risk groups, with fewer patients in the low IPI group (24.7% vs 32.4%; $P = .004$) and more patients in the high IPI group (27.0% vs 19.5%; $P = .002$). During a median 41-month observation period, the 3-year TTP, PFS, and OS in our study were 69.3% (95% confidence interval [CI], 64.1-74.0), 60.3% (95% CI, 55.0-65.1), and 72.0% (95% CI, 67.1-76.4), respectively, which were lower than those of the original study (79.7% [95% CI, 77.4-82.0], 74.5% [95% CI, 72.1-77.0], and 81.8% [95% CI, 79.7-84.0], respectively). Reflecting age–associated vulnerabilities, 41 patients (10.6%) died without lymphoma progression, with 12 (3.1%) from infections and 10 (2.6%) from other cancers.

Remarkably, after ranking individual IMPI scores, almost half of the patients (n = 195 [50.6%]) were reclassified within IMPI categories compared with the original IPI classifications (Figure 1A). Among these, 100 (25.9%) moved to the higher-risk groups and 95 (24.7%) to the lower-risk groups. The overall agreement between IMPI and $IMPI_{median}$ was 90.1%, with the highest concordance observed in the low-risk IMPI groups at 100%, whereas the high-intermediate group showed a relatively high discordance rate of 20.7% (Figure 1B). Of note, IMPI-based allocation did not clearly improve outcome discrimination (TTP in Figure 1C; PFS and OS in supplemental Figure 3), except in the IPI low-intermediate group, which exhibited a trend toward worse TTP in patients with upgraded IMPI (3-year TTP, 70.4%; 95% CI, 55.0-90.0; $P = .088$). Using the $IMPI_{median}$, the discrimination power in the IPI low-intermediate group increased, with the trend reaching statistical significance (3-year TTP: IMPI-downgraded, 90.8% [95% CI, 79.3-100]; IMPI-not reclassified, 90.4% [95% CI, 80.5-100]; and IMPI-upgraded, 65.7% [95% CI, 50.4-85.5]; $P = .006$) (Figure 1C; other survival curves by $IMPI_{median}$ in supplemental Figure 4). The comparison of background characteristics between the IMPI-downgraded and -upgraded populations revealed that none with advanced disease were included in the former group (vs 43.3% in the latter; $P < .001$), suggesting that IMPI more efficiently identified localized disease within the IPI low-intermediate group and contributed to improved stratification (supplemental Table 1). Importantly, in contrast to the original study, the IMPI-based 3-group separation failed to identify a high-risk group with poorer outcomes than the high-risk groups

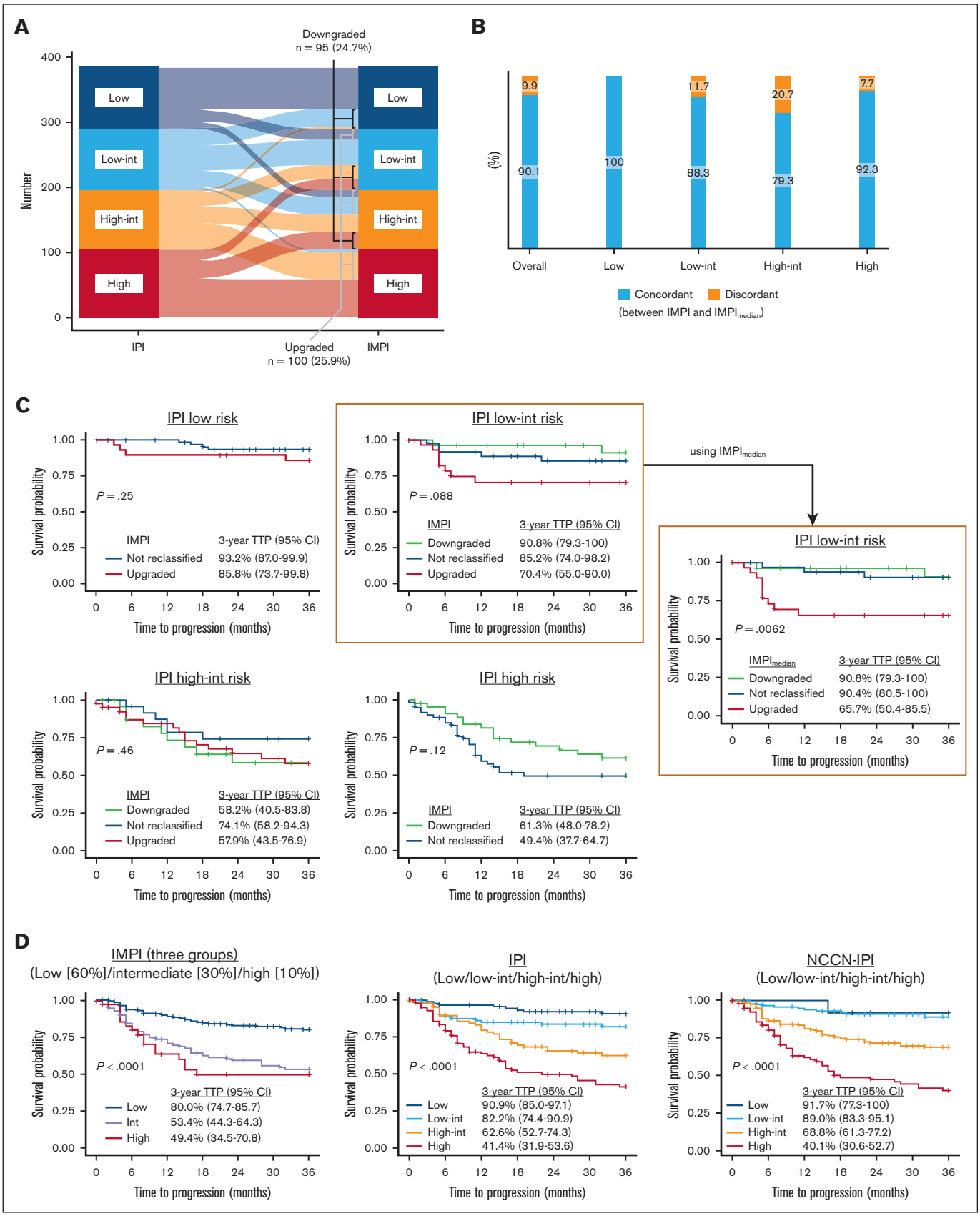

Figure 1.

Figure 1. Reclassification of patients from IPI to IMPI and its prognostic impact. (A) An alluvial diagram depicting the transition from IPI to IMPI categories. (B) The concordance of IMPI classification when using different TMTV cutoffs: the original study's 307.9 mL vs this study's median value of 198.2 mL. (C) Kaplan-Meier estimates of TTP among the 4 IPI categories stratified by IMPI reclassification. The TTP of IPI low-int patients reclassified by $IMPI_{median}$ is also shown. (D) Kaplan-Meier estimates of TTP according to IMPI 3 (low-intermediate-high) categorization, IPI, and NCCN-IPI. High-int, high-intermediate; Low-int, low-intermediate.

defined by the IPI and NCCN-IPI (3-year TTP: IMPI high-risk, 49.4% [95% CI, 34.5-70.8]; IPI high-risk, 41.4% [95% CI, 31.9-53.6]; and NCCN-IPI high-risk, 40.1% [95% CI, 30.6-52.7]; Figure 1D; PFS and OS in supplemental Figure 5). When categorized into 4 groups to match the sizes, the survival curves for IMPI were similar to those of IPI and NCCN-IPI (supplemental Figure 6).

Finally, we compared predictive performance and model fitness for the 3 indices listed as categorical and continuous variables in supplemental Table 2. The IMPI-based categorization yielded a lower c-index and a higher AIC than both IPI and NCCN-IPI, regardless of variable types, except for the c-index of $IMPI_{median}$ (the same size as IPI) for TTP (c-index, 0.698). In contrast, NCCN-IPI, when analyzed as a continuous variable, exhibited the best model performance (TTP: AIC, 1119.93; c-index, 0.714; PFS: AIC, 1583.73; c-index, 0.691; OS: AIC, 1116.84; c-index, 0.694).

The 2 clinical trials have assessed the predictive abilities of IMPI. One, focusing on risk-adapted immunochemotherapy, found IMPI overestimated event rates,[5] whereas the other, involving patients with relapsed/refractory DLBCL treated with locastuximab tesirine, showed IMPI performed worse than TMTV alone.[6] Similarly, our study demonstrated that IMPI did not outperform the conventional IPI and NCCN-IPI systems in patients with newly diagnosed DLBCL treated with R-CHOP or R-CHOP–like regimen. To our knowledge, this is the first real-world validation study of IMPI, except for 1 report in the specific context of chimeric antigen receptor T-cell therapy.[7]

The original study appeared to underrepresent populations outside of the Positron Emission Tomography ReAnalysis consortium. Our study included significantly more patients who were older, frail, and relatively less advanced. This cohort profile aligns with the recent large real-world data sets from Japan (n = 1050; 83% older, 18% frail, and 55% stage III-IV)[8] and Vancouver (n = 1149; 72% older, 43% frail, and 59% stage III-IV)[9] (supplemental Table 3). Moreover, median TMTV values, which form the regression formula of IMPI, are highly influenced by the extent of high tumor burden in the analyzed cohorts. Our median value of 198 mL is approximately two-thirds of the original study but close to that reported in a recent Japanese study using the same SUV 4.0 segmentation threshold (236 mL for the training cohort and 167 mL for the validation cohort).[10] We used cohort-specific median TMTV values to establish $IMPI_{median}$, with an acceptable concordance rate of 90.1%. However, even using $IMPI_{median}$, only patients in the IPI low-intermediate group showed significantly better outcome discrimination with this TMTV-based system. Regarding c-index and AIC, almost all IMPI-based categorizations provided inferior values to that of IPI and NCCN-IPI. These results suggest that outcomes of real-world patients may be more influenced by patient factors such as age and performance status than tumor burden,[11] which potentially reduces the applicability of the IMPI system.

Taken together, the IMPI model is not as robust as anticipated in the original study. The IMPI did not outperform IPI and NCCN-IPI, being at best equivalent. The main limitations of our study are its retrospective nature and relatively small sample size when stratified into 4 IPI groups, despite the high methodological consistency with the original study. Our cohort profiles, which include a higher number of individuals who were older and frail potentially leading to increased nonlymphoma mortality, as well as lower tumor burdens and a greater proportion of IPI high-risk groups, may affect the validity of this comparative study. Further large-scale validation is needed to potentially develop modifications for real-world populations before widespread IMPI application in clinical practice.

Acknowledgments: The authors thank Hajime Senjo (Department of Hematology, Hokkaido University) and Kenji Hirata (Department of Nuclear Medicine, Hokkaido University) for their assistance with the Metavol software. The authors also thank Editage for providing excellent assistance with the English language editing.

Contribution: D.I. and K.M. designed the study, interpreted the data, performed statistical analysis, provided patient care, and wrote the manuscript; D.I. calculated the quantitative radiomic features; Y.M. interpreted the images. M.O., A.U., R.T., K.N., and M.T. provided patient care; and all authors critically reviewed and approved the manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

ORCID profiles: D.I., 0000-0002-7398-4616; M.O., 0000-0002-4907-4647; K.N., 0000-0002-6504-9046; K.M., 0000-0002-8669-9865.

Correspondence: Daisuke Ikeda, Division of Hematology/Oncology, Department of Medicine, Kameda Medical Center, 929 Higashi-chou, Kamogawa-shi, Chiba 296-8602, Japan; email: dskikd.2409@gmail.com.

References

1. Mikhaeel NG, Heymans MW, Eertink JJ, et al. Proposed New Dynamic Prognostic Index for diffuse large B-cell lymphoma: International Metabolic Prognostic Index. *J Clin Oncol*. 2022;40(21):2352-2360.
2. Eertink JJ, Zwezerijnen GJC, Heymans MW, et al. Baseline PET radiomics outperforms the IPI risk score for prediction of outcome in diffuse large B-cell lymphoma. *Blood*. 2023;141(25):3055-3064.
3. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med*. 2006;25(1):127-141.
4. Zhou Z, Sehn LH, Rademaker AW, et al. An enhanced International Prognostic Index (NCCN-IPI) for patients with diffuse large B-cell lymphoma treated in the rituximab era. *Blood*. 2014;123(6):837-842.
5. Michaud L, Bantilan K, Mauguen A, Moskowitz CH, Zelenetz AD, Schoder H. Prognostic value of (18)F-FDG PET/CT in diffuse large B-cell lymphoma treated with a risk-adapted immunochemotherapy regimen. *J Nucl Med*. 2023;64(4):536-541.
6. Alderuccio JP, Reis IM, Hamadani M, et al. PET/CT-biomarkers enable risk stratification of patients with relapsed/refractory diffuse large B-cell lymphoma enrolled in the LOTIS-2 clinical trial. *Clin Cancer Res*. 2024;30(1):139-149.

7. Winkelmann M, Blumenberg V, Rejeski K, et al. Prognostic value of the International Metabolic Prognostic Index for lymphoma patients receiving chimeric antigen receptor T-cell therapy. *Eur J Nucl Med Mol Imaging*. 2023;50(5):1406-1413.

8. Urata T, Naoi Y, Jiang A, et al. Distribution and clinical impact of molecular subtypes with Dark Zone signature of DLBCL in a Japanese real-world study. *Blood Adv*. 2023;7(24):7459-7470.

9. Alduaij W, Collinge B, Ben-Neriah S, et al. Molecular determinants of clinical outcomes in a real-world diffuse large B-cell lymphoma population. *Blood*. 2023;141(20):2493-2507.

10. Senjo H, Kanaya M, Izumiyama K, et al. Serum level of soluble interleukin-2 receptor is positively correlated with metabolic tumor volume on (18) F-FDG PET/CT in newly diagnosed patients with diffuse large B-cell lymphoma. *Cancer Med*. 2019; 8(3):953-962.

11. Adams HJ, de Klerk JM, Fijnheer R, et al. Prognostic superiority of the National Comprehensive Cancer Network International Prognostic Index over pretreatment whole-body volumetric-metabolic FDG-PET/CT metrics in diffuse large B-cell lymphoma. *Eur J Haematol*. 2015;94(6):532-539.