

ORIGINAL RESEARCH

Prognostic impact of six versus eight cycles of standard regimen in patients with diffuse large B-cell lymphoma: propensity score-matching analysis

K. Oiwa^{1,2†}, K. Fujita^{1,3†}, S. Lee^{1,3*}, T. Morishita⁴, H. Tsukasaki^{1,5}, E. Negoro¹, T. Hara³, H. Tsurumi³, T. Ueda¹ & T. Yamauchi¹

¹Department of Hematology and Oncology, University of Fukui, Fukui; ²Department of Hematology and Oncology, Nagoya City University, Aichi; ³Department of Hematology, Matsunami General Hospital, Gifu; ⁴Department of Healthcare Economics and Quality Management, Graduate School of Medicine, Kyoto University, Yoshida Konoe-cho, Kyoto; ⁵Department of Hematology, Fukui Red Cross Hospital, Fukui, Japan



Available online xxx

Background: R-CHOP-21 has been the standard treatment for diffuse large B-cell lymphoma (DLBCL), but there is a paucity of evidence focusing on the number of cycles of regimens.

Patients and methods: We conducted a retrospective study to compare the effectiveness of six cycles of standard regimens versus eight cycles for overall survival (OS) in DLBCL patients using propensity score matching, in consideration of relative dose intensity (RDI).

Results: A total of 685 patients with newly diagnosed DLBCL were identified in three institutions from 2007 to 2017. Patients treated using six cycles of standard regimens were matched by propensity scores with those treated using eight cycles. A 1 : 1 propensity score matching yielded 138 patient pairs. Eight cycles did not significantly improve OS in the conventional Cox proportional hazards model (hazard ratio 0.849, 95% confidence interval 0.453-1.588, $P = 0.608$). Restricted cubic spline Cox models for OS confirmed that the effect of the number of cycles was not modified by total average RDI, the International Prognostic Index, and age. Occurrence of adverse events did not differ between six and eight cycles.

Conclusion: Even considering the impact of RDI, six cycles of the initial standard regimen for DLBCL is not inferior to eight cycles.

Key words: diffuse large B-cell lymphoma, relative dose intensity, number of chemotherapy cycles, propensity score matching, restricted cubic spline

INTRODUCTION

For several decades, CHOP [cyclophosphamide (CPA), doxorubicin (DXR), vincristine (VCR), and prednisone (PSL)] therapy combined with rituximab (R-) has been considered the standard regimen for diffuse large B-cell lymphoma (DLBCL).¹ Conventionally, six cycles of R-CHOP-21 have frequently been used in clinical practice, but data supporting the optimal number of treatment cycles remain limited.^{2,3} While no randomized trials have focused on the number of cycles of R-CHOP-21, one prospective trial restricted to elderly patients showed the non-inferiority of six cycles of R-CHOP-14 versus eight cycles.²

Maintaining a higher relative dose intensity (RDI) of chemotherapeutic drugs plays an important role in attempts

to achieve better outcomes in the management of aggressive lymphomas, including DLBCL.⁴⁻⁷ Based on the calculation method of the RDI, the number of treatment cycles has a significant impact on total average RDI (tARDI) during the entire treatment period.^{4,6,7} Recently, one retrospective study using matched sample analysis addressed the possibility that six cycles of R-CHOP-21 is not inferior to eight cycles for DLBCL.³ However, the RDI of standard regimens had not been evaluated in the reported analyses.^{3,4}

We therefore designed the present study to verify the prognostic impact of differences in the number of cycles of a standard regimen (six versus eight cycles), taking RDI into consideration. To evaluate this clinical issue, we used propensity score matching and Cox hazards modeling with restricted cubic spline (RCS). Furthermore, subgroup analyses were carried out to assess whether any populations would benefit from either six or eight cycles of standard regimen.

PATIENTS AND METHODS

Study population

This multicenter, retrospective analysis examined a cohort from three tertiary institutions in Japan: University of Fukui

*Correspondence to: Dr Shin Lee, Department of Hematology, Matsunami General Hospital, Dendai 185-1 Kasamatsu-cho, Hashima-gun, Gifu 501-6062, Japan. Tel: +81-583-88-0111; Fax: +81-583-88-4711
E-mail: leesin.581020@gmail.com (S. Lee).

† These authors contributed equally to this work.

2059-7029/© 2021 The Authors. Published by Elsevier Ltd on behalf of European Society for Medical Oncology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Hospital, the Japanese Red Cross Fukui Hospital, and Fukui Prefectural Hospital. Data were identified using medical records from the participating institutions. Patients with newly diagnosed DLBCL, ≥ 18 years old, were identified from the three participating institutions from 2007 to 2017. To identify patients receiving six or eight cycles of standard regimen alone, patients who received one to five or seven cycles of standard regimen, non-standard regimen, or radiotherapy alone were excluded. Patients with central nervous system involvement, post-transplant or methotrexate-associated lymphoproliferative disorders, or transformed DLBCL were also excluded. Histologic diagnoses were made according to the Revised European American Lymphoma classification and the World Health Organization classification.^{8,9} Patient age, sex, anthropometric data (height and weight), performance status (PS), number of extranodal sites, stage, elevated lactate dehydrogenase, serum albumin (Alb), serum soluble interleukin-2 receptor (sIL-2R),^{10,11} International Prognostic Index (IPI), B symptoms (fevers, night sweats, or unintentional weight loss), and bulky mass (diameter >7.5 cm) were collected as baseline demographic data.^{12,13} Comorbidity and nutrition status at diagnosis were assessed by the Charlson Comorbidity Index (CCI) and Geriatric Nutritional Risk Index (GNRI).¹⁴⁻¹⁷ The Geriatric 8 (G8) score was used to assess frailty in the elderly patients aged ≥ 65 years.^{18,19} This study was conducted in accordance with the principles of the Declaration of Helsinki, and all protocols were approved by the appropriate institutional review boards. The need to obtain written, informed consent was waived, since this study used retrospective data obtained only from hospital medical records.

Treatment definitions and calculation of RDI

The standard regimens in this study included the combination of rituximab 375 mg/m² and CHOP [CPA 750 mg/m² intravenously on day 1, DXR 50 mg/m² intravenously on day 1, VCR 1.4 mg/m² (maximum 2.0 mg/body) intravenously on day 1, PSL 100 mg/body orally or intravenously on days 1-5; (R-) CHOP], or THP-COP [CPA, tetrahydropyranil adriamycin (THP), VCR, PSL; (R-) THP-COP] every 3 weeks.^{1,20} The THP-COP regimen was the same as CHOP, including the doses, except that THP replaced DXR.

RDI was assessed throughout the entire treatment period as tARDI in this study. Details of the method to calculate tARDI are described in the [Supplementary Materials](https://doi.org/10.1016/j.esmoop.2021.100210), available at <https://doi.org/10.1016/j.esmoop.2021.100210>. Six cycles of standard regimen without any dose reduction or administration delay was defined as a tARDI value of 100%.^{4,5} Based on this calculation method, tARDI was strongly influenced by the number of cycles of treatment. In cases treated with eight cycles of standard regimens, tARDI could thus exceed 100%.

Outcome measures

The primary outcome in our study was overall survival (OS). Prognostic factors generally thought to be associated with

survival were also evaluated. Moreover, subgroup analysis by IPI was carried out. OS was calculated from the date of final diagnosis to the date of death from any cause, or the last recorded follow-up visit. In other words, the date of final diagnosis represented the date they entered the cohort, or time zero. Patients were followed up until the end of the assessment (31 October 2018) or death, whichever came first.

Common Terminology Criteria for Adverse Events version 4.0 were used to document treatment-related toxicities.²¹ We defined grade ≥ 3 treatment-related toxicities as severe adverse events (AEs) in this study.

Statistical analysis

We defined two groups for analysis: six and eight cycles of standard regimens. We used propensity score matching to control for observed confounding factors that might influence the number of chemotherapy cycles, and detailed statistical methods are described in the [Supplementary Materials](https://doi.org/10.1016/j.esmoop.2021.100210), available at <https://doi.org/10.1016/j.esmoop.2021.100210>. Sensitivity analyses were also verified using the original dataset.

Patient characteristics were examined for six versus eight chemotherapy cycles, in addition to survivors versus non-survivors. Continuous variables are expressed as median values and ranges, and groups were compared using the Mann–Whitney *U* test. Categorical variables are expressed as numbers and percentages, and groups were compared using the chi-square test or Fisher's exact test, as appropriate. Survival curves for each group were estimated using Kaplan–Meier methods, with comparison between groups carried out using log-rank testing. Hazard ratios (HRs) and 95% confidence intervals (CIs) between groups were determined using Cox proportional hazards modeling adjusted for sex, IPI, bulky mass, CCI, GNRI, tARDI and eight chemotherapy cycles. Non-linear regression with RCS of three knots was used to examine for the presence of non-linear relationships between OS and tARDI and between OS and age. Interaction analysis for age and tARDI was carried out in six versus eight cycles. RCS for each IPI score was carried out as a subgroup analysis.

We also conducted a sensitivity analysis controlling for the immortal bias in observational cohort studies. To compare OS considering the time-varying effect of receiving chemotherapy, time-varying Cox proportional hazards regression analysis was carried out. In order to further evaluate the consistency and robustness of main analysis, the relationship between the number of cycles of standard regimens and the outcome variable, the chemotherapy of standard regimens was included as a time-dependent incidence variable, and encoded as 0 for periods without chemotherapy and as 1 for periods with chemotherapy. Other baseline covariates included in the time-varying Cox regression model were sex, IPI, bulky mass, CCI, GNRI, tARDI and eight chemotherapy cycles. The time-varying Kaplan–Meier curves were created according to the methods described by Simon and Makuch.^{22,23}

Table 1. Patient characteristics at diagnosis after propensity score matching

	All patients (N = 276)	Six cycles of standard regimens (n = 138)	Eight cycles of standard regimens (n = 138)	P value	SMD	Survivor (n = 203)	Non-survivor (n = 73)	P value
Age, years, median (range)	72 (27-90)	72 (33-90)	72 (27-87)	0.345		72 (27-90)	72 (33-89)	0.482
≥70, n (%)	155 (56.2)	75 (54.3)	80 (58.0)	0.628	0.073	114 (56.2)	41 (56.2)	0.999
Male, n (%)	144 (52.2)	71 (51.4)	73 (52.9)	0.904	0.029	101 (49.8)	43 (58.9)	0.219
ECOG PS ≥2, n (%)	59 (21.4)	33 (23.9)	26 (18.8)	0.378		32 (15.8)	27 (37.0)	<0.001
Extranodal sites ≥2, n (%)	96 (34.8)	45 (32.6)	51 (37.0)	0.528		69 (34.0)	27 (37.0)	0.669
Ann Arbor stage III/IV, n (%)	187 (67.8)	90 (65.2)	97 (70.3)	0.440		130 (64.0)	57 (78.1)	0.029
Elevated LDH (>ULN), n (%)	181 (65.6)	86 (62.3)	95 (68.8)	0.311		123 (60.6)	58 (79.5)	0.004
Serum albumin (g/dl), median (range)	3.6 (1.1-5.1)	3.5 (1.4-5.1)	3.6 (1.1-4.9)	0.648		3.6 (1.1-5.1)	3.4 (1.6-4.9)	0.006
IPI, n (%)								
Low/low intermediate (0-2)	117 (42.4)	61 (44.2)	56 (40.6)	0.626	0.073	96 (47.3)	21 (28.8)	0.006
High intermediate/high (≥3)	159 (57.6)	77 (55.8)	82 (59.4)			107 (52.7)	52 (71.2)	
NCCN-IPI, n (%)								
Low/low intermediate (0-3)	96 (34.8)	49 (35.5)	47 (34.1)	0.899	0.030	81 (39.9)	15 (20.5)	0.003
High intermediate/high (≥4)	180 (65.2)	89 (64.5)	91 (65.1)			122 (60.1)	58 (79.5)	
Bulky mass, n (%)	52 (18.8)	24 (17.4)	28 (20.3)	0.645	0.074	34 (16.7)	18 (24.7)	0.163
B symptoms, n (%)	91 (33.0)	43 (31.2)	48 (34.8)	0.609		61 (30.0)	30 (41.1)	0.110
Uric acid (mg/dl), median (range)	5.3 (0.2-23.0)	5.4 (0.2-13.7)	5.0 (1.1-23.0)	0.327		5.1 (0.2-23.0)	6.0 (0.7-13.7)	0.052
≥7.5 mg/dl	37 (13.4)	18 (13.0)	19 (13.8)	0.999	0.021	23 (11.3)	14 (19.2)	0.109
sIL-2R, median (range)	1441 (168-38 400)	1343 (168-38 400)	1478 (260-25 400)	0.943		1200 (168-25 400)	3021 (316-38 400)	<0.001
≥Median	145 (52.5)	70 (50.7)	75 (54.3)	0.630	0.073	95 (46.8)	50 (68.5)	0.002
CCI, n (%)								
Low/medium (0-2)	234 (84.8)	116 (84.1)	118 (85.5)	0.867	0.040	173 (85.2)	61 (83.6)	0.708
High/very high (≥3)	42 (15.2)	22 (15.9)	20 (14.5)			30 (14.8)	12 (16.4)	
GNRI, n (%)								
No-risk group (>98)	104 (37.7)	52 (37.7)	52 (37.7)	0.999	<0.001	66 (32.5)	38 (52.1)	0.005
Risk group (≤98)	172 (62.3)	86 (62.3)	86 (62.3)			137 (67.5)	35 (47.9)	
Total ARDI	98.8 (119.2-142.6)	91.6 (19.2-106.4)	123.1 (61.2-142.6)	<0.001		99.3 (35.4-142.6)	92.4 (19.2-138.0)	0.067
ASCT	17 (6.2)	14 (10.1)	3 (2.2)	0.010		8 (3.9)	9 (12.3)	0.020
Up-front ASCT	12 (4.3)	12 (8.7)	0 (0.0)	<0.001		6 (3.0)	6 (8.2)	0.088
ASCT as salvage therapy	5 (1.8)	2 (1.4)	3 (2.2)	0.999		2 (1.0)	3 (4.1)	0.117

ASCT, autologous stem cell transplantation; CCI, Charlson Comorbidity Index; ECOG PS, Eastern Cooperative Oncology Group performance status; GNRI, Geriatric Nutritional Risk Index; IPI, International Prognostic Index; LDH, lactate dehydrogenase; NCCN, National Comprehensive Cancer Network; sIL-2R, soluble interleukin-2 receptor; SMD, standardized mean difference; ULN, upper limit of normal.

We carried out a multivariate logistic regression analysis using the original dataset to clarify the factors related to severe AEs. Factors examined included sex, the IPI score, serum Alb, bulky mass, tARDI, CCI score and the number of cycles of standard regimens (six versus eight cycles). All *P* values were two-sided. *P* values <0.05 and *P* values for interaction <0.10 were considered significant. Data analyses were carried out using R version 3.6.2 or EZR version 1.37, which is a graphical user interface for R.^{24,25}

RESULTS

Patient characteristics and propensity score matching

A total of 685 patients were diagnosed with *de novo* DLBCL at three participating tertiary institutions. Of these, 67 patients were excluded based on the exclusion criteria relating to disease status at diagnosis, and 274 patients were excluded because of the exclusion criteria relating to type of regimen or number of courses of chemotherapy. The remaining 371 patients who received treatment with six or eight cycles of standard regimens were enrolled into this study as the original dataset. Of these, 191 patients were treated with six cycles, and 180 patients were treated with eight cycles (Supplementary Figure S1, available at <https://doi.org/10.1016/j.esmooop.2021.100210>). Patients with the

germinal center (GCB) were 35, patients with activated B-cell lymphoma (ABC) were 72, and the remaining 264 patients were unknown. Patient characteristics in the original dataset before matching are shown in Supplementary Table S1, available at <https://doi.org/10.1016/j.esmooop.2021.100210>. Patient characteristics between included patients and excluded patients are presented in Supplementary Table S2, available at <https://doi.org/10.1016/j.esmooop.2021.100210>. There were significant differences in the patients' background, including age, sex, PS, serum Alb level, the CCI, and the GNRI, between the patients who could not complete six or eight cycles of the standard regimens (371 patients) and those who completed six or eight cycles (189 patients).

Table 1 shows that the distribution of baseline covariates was adequately balanced in the matched dataset. Standardized mean difference in the matched dataset in this cohort was <0.1, suggesting a negligible difference between the six-cycle and eight-cycle groups. Histograms of propensity scores in both datasets before and after matching showed that the distributions of propensity scores between the six-cycle and eight-cycle groups almost completely overlapped after matching (Supplementary Figure S2A and B, available at <https://doi.org/10.1016/j.esmooop.2021.100210>). The c-statistic for propensity scores was

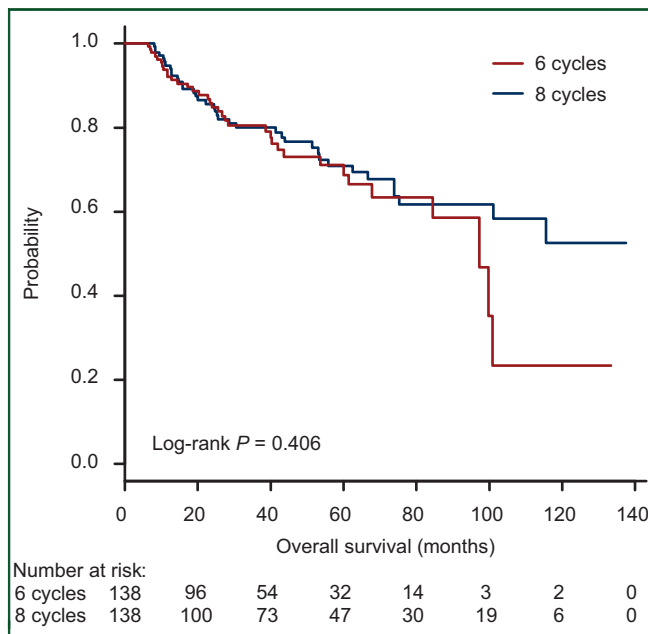


Figure 1. Kaplan–Meier survival curve of overall survival using the matched dataset according to the number of cycles of regimens.

0.63 (95% CI 0.58-0.96) (Supplementary Figure S3, available at <https://doi.org/10.1016/j.esmooop.2021.100210>).

In the matched dataset, patient characteristics at diagnosis and comparing six cycles and eight cycles, and survivor and non-survivor groups are summarized in Table 1. Median age at diagnosis was 72 years (range, 27-90 years) and 71.0% (196/276) of the patients were aged ≥65 years. There was no significant difference in the G8 score between the two groups (six versus eight cycles) only in the elderly population aged ≥65 years ($P = 0.384$). The six-cycle group showed significantly higher tARDI than the eight-cycle group. Patients who underwent high-dose chemotherapy (HDT) with autologous stem cell transplantation (ASCT) were significantly more frequent in the six-cycle group than in the eight-cycle group, but less frequent in the survivor group than in the non-survivor group.

Overall survival outcomes

Kaplan–Meier survival curves for the six-cycle and eight-cycle groups using the matched dataset are shown in Figure 1. No significant difference in OS was seen between groups depending on the number of cycles of standard regimens (log-rank $P = 0.406$). The same result was also demonstrated in the Kaplan–Meier survival curves using the original dataset (Supplementary Figure S4, available at <https://doi.org/10.1016/j.esmooop.2021.100210>).

Predictors of OS

Median follow-up in the matched dataset was 34.3 months (range, 4.4-137.5 months), and 73 patients (26.4%) died (35 patients in the six-cycle group versus 38 patients in the eight-cycle group). No patient died of AEs during the treatment period. Multivariate Cox hazards proportional models for OS were created to evaluate the impact of

Table 2. Multivariate Cox proportional hazards analysis of clinical factors significantly associated with overall survival

	HR (95% CIs)	P value
Male	1.751 (1.060-2.893)	0.029
IPI, score	1.341 (1.091-1.648)	0.005
Bulky mass	1.520 (0.866-2.670)	0.145
CCI, score	1.172 (1.014-1.355)	0.032
GNRI, score	0.985 (0.967-1.003)	0.096
tARDI, %	0.998 (0.986-1.010)	0.730
Eight cycles of standard regimens	0.849 (0.453-1.588)	0.608

CCI, Charlson Comorbidity Index; CIs, confidence intervals; GNRI, Geriatric Nutritional Risk Index; HR, hazard ratio; IPI, International Prognostic Index; tARDI, total average relative dose intensity.

prognostic covariates (Table 2). Male sex, IPI, and CCI were the independent predictors for all-cause mortality. The difference in number of cycles of standard regimens did not show any significant influence on survival (HR 0.849, 95% CI 0.453-1.588, $P = 0.608$). There was no interaction of the subtype (GCB, ABC, or unknown) on the relationship between the number of cycles of standard regimens and survival in post hoc analysis (P for interaction = 0.185).

The non-significance of six cycles versus eight cycles persisted in the sensitivity analysis including the time-varying Cox regression analysis and time-varying Kaplan–Meier curves (Supplementary Table S3 and Figure S5, available at <https://doi.org/10.1016/j.esmooop.2021.100210>). Results of the time-varying Cox regression analysis remained similar to those of the previous main analysis (HR 1.538, 95% CI 0.974-2.498, $P = 0.082$ for the effect of eight cycles of standard regimens) (Supplementary Table S3, available at <https://doi.org/10.1016/j.esmooop.2021.100210>). The difference in number of cycles of standard regimens did not show any significant influence on survival in the time-varying Kaplan–Meier curves (non-adjusted HR 1.052, 95% CI 0.725-1.527, $P = 0.788$) (Supplementary Figure S5, available at <https://doi.org/10.1016/j.esmooop.2021.100210>).

There was no significant difference in interim treatment response sIL-2 values after four cycles of standard regimens between the two groups (six versus eight cycles) both in the original dataset (median 504, range 209-1950 in the six-cycle group versus median 475, range 54.6-2790.0 in the eight-cycle group, $P = 0.390$) and the matched dataset (median 506, range 214-1960.0 in the six-cycle group versus median 521, range 54.6-2790.0 in the eight-cycle group, $P = 0.897$).

Effect modification of six cycles versus eight cycles with impact of tARDI on OS

To assess whether the number of cycles of standard regimens had an interaction effect with the impact of tARDI on OS, we created an RCS-Cox model in the binary groups (six versus eight cycles) (Figure 2). For any tARDI, when the tARDI was the same, no significant difference in mortality risk was seen between the six-cycle and eight-cycle groups. The number of cycles (six versus eight cycles) of standard

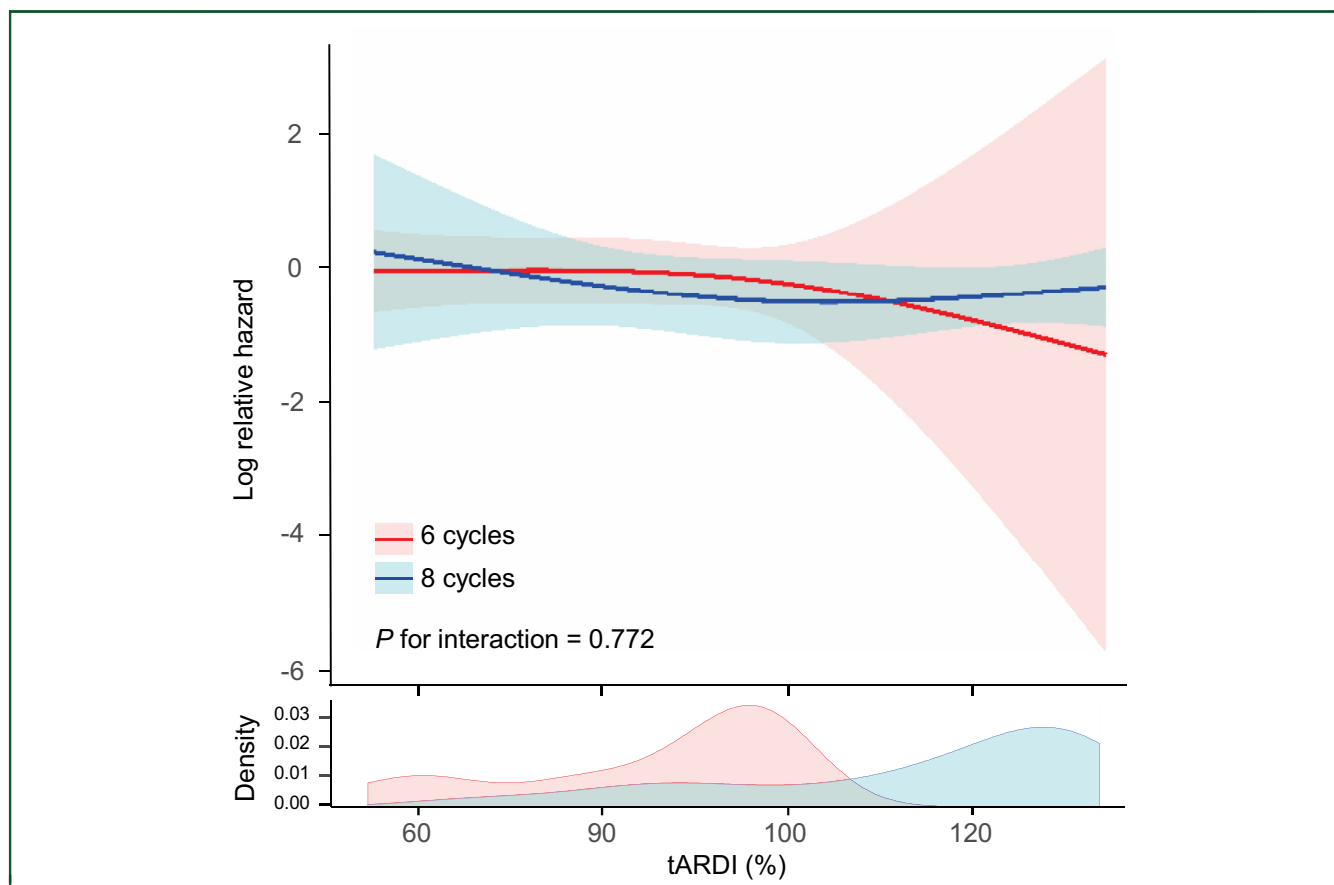


Figure 2. Covariate-adjusted Cox hazards model with restricted cubic spline with three knots showing the association between tARDI and overall survival according to the number of cycles of regimens using the matched dataset.

Solid line represents the log hazard ratio. Shaded area is the 95% confidence interval. tARDI, total average relative dose intensity.

regimens did not show an interaction with the impact of tARDI on mortality risk. Lack of interaction of the number of cycles (six versus eight cycles) with the relationship between tARDI and OS was also verified in the original dataset (Supplementary Figure S6, available at <https://doi.org/10.1016/j.esmoop.2021.100210>).

Effect modification of six cycles versus eight cycles with impact of age on OS

We also created an RCS-Cox model in binary groups (six cycles versus eight cycles) to assess the interaction of the number of cycles of standard regimens with the relationship between age and OS (Figure 3). No difference in mortality risk was seen between six cycles and eight cycles at any age. The same result was obtained in the analysis using the original dataset (Supplementary Figure S7, available at <https://doi.org/10.1016/j.esmoop.2021.100210>).

Effect modification of IPI for relationship between tARDI and OS

Supplementary Figure S8, available at <https://doi.org/10.1016/j.esmoop.2021.100210>, shows the interaction of the number of cycles (six cycles versus eight cycles) with the relationship between tARDI and OS in each IPI category based on an RCS-Cox model. The effect of the number of

cycles of regimen (six cycles versus eight cycles) on the relationship between tARDI and OS was not modified by IPI. Subgroup analyses using the original dataset yielded the same result (Supplementary Figure S9, available at <https://doi.org/10.1016/j.esmoop.2021.100210>). In both of the binary groups, depending on IPI (<3 versus ≥3), the number of cycles of regimens did not interact with the effect of tARDI on mortality risk (Supplementary Figure S10A, available at <https://doi.org/10.1016/j.esmoop.2021.100210>). Sensitivity analysis using the original dataset showed the same result (Supplementary Figure S10B, available at <https://doi.org/10.1016/j.esmoop.2021.100210>).

Adverse events

Supplementary Table S3, available at <https://doi.org/10.1016/j.esmoop.2021.100210>, shows the occurrence of severe non-hematological toxicity and febrile neutropenia in the matched dataset and original dataset. The frequency of severe AEs did not differ significantly between the six-cycle and eight-cycle groups in both datasets before and after matching (Supplementary Table S4, available at <https://doi.org/10.1016/j.esmoop.2021.100210>). The multivariate logistic regression analysis using the original dataset showed that the number of cycles of regimens (six versus eight cycles) did not significantly influence the risk of severe AEs

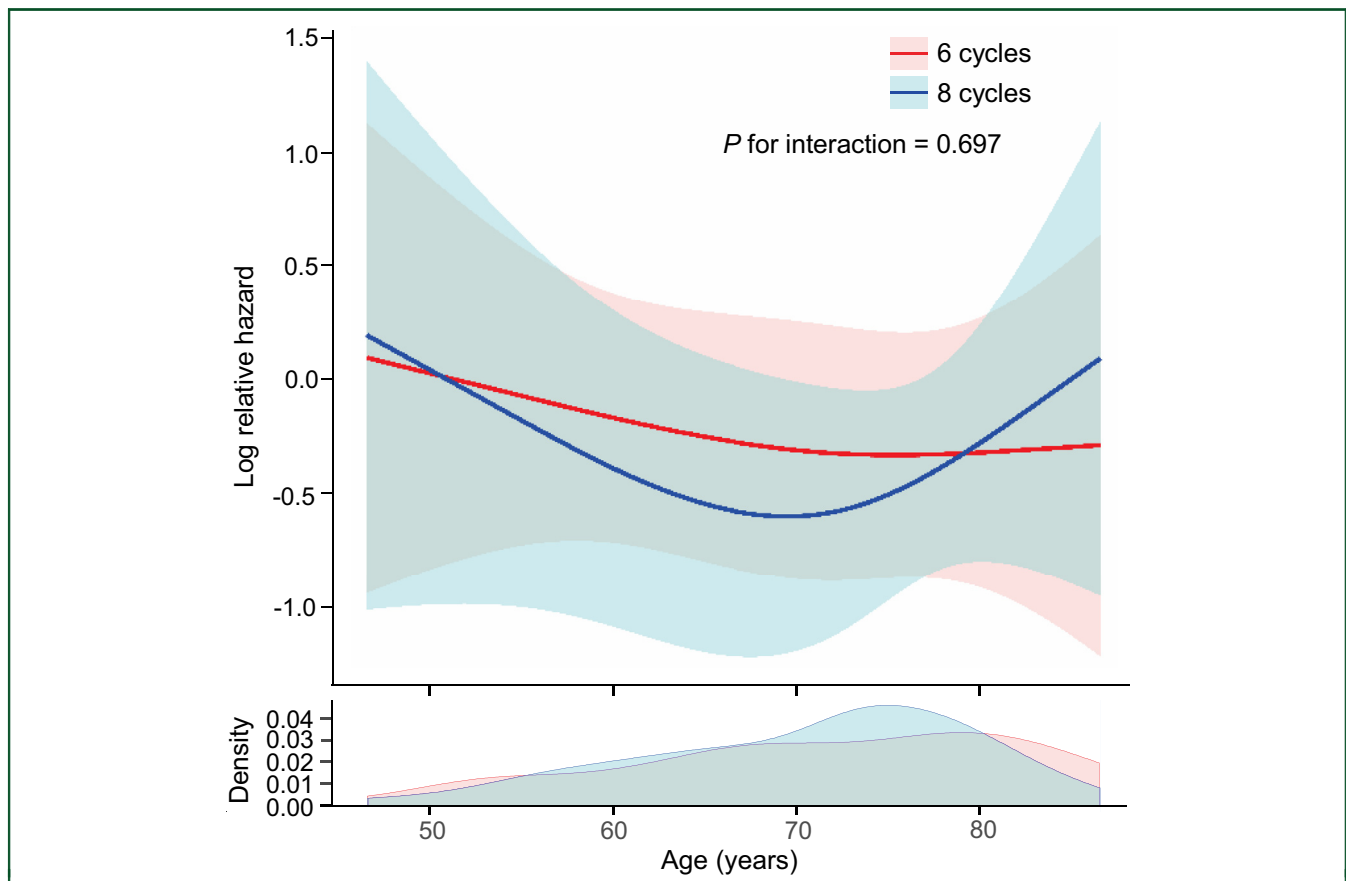


Figure 3. Covariate-adjusted Cox hazards model with restricted cubic spline with three knots showing the association between age and overall survival according to the number of cycles of regimens using the matched dataset.

Solid line represents the log hazard ratio. Shaded area is the 95% confidence interval.

(Supplementary Table S5, available at <https://doi.org/10.1016/j.esmooop.2021.100210>).

DISCUSSION

Six cycles of standard regimens for DLBCL did not result in inferior prognosis compared to eight cycles, even taking RDI into consideration. The non-inferiority of six cycles compared with eight cycles was also demonstrated at any age, and in any IPI group. These results were assessed by the RCS-Cox model that can provide continuous visual information on relationships between variables. The robustness of these results was guaranteed by sensitivity analyses using the original dataset.

To the best of our knowledge, our study represents the first report to examine whether the number of treatment cycles (six versus eight cycles) interacts with the relationship between RDI and OS. The number of treatment cycles (six versus eight cycles) was shown to have no significant impact on OS in either conventional or RCS-Cox models. The robustness of this result was validated in sensitivity analyses. RDI has a significant impact on the prognosis of DLBCL, not only in terms of average RDI per cycle, but also in terms of tARDI over the entire treatment period.^{5,7} Based on the calculation of RDI, as the number of cycles increases, tARDI increases proportionally.⁴ A previous report

demonstrated the non-inferiority of six cycles of R-CHOP-21 to eight cycles for DLBCL, but the interaction of RDI was not validated.³ Validation including RDI is necessary, because differences in the number of treatment cycles results in non-negligible differences in tARDI.

In our study, tARDI was significantly higher in the eight-cycle group than in the six-cycle group. However, eight-cycle chemotherapy was not associated with improved survival compared to six cycles. Several reports have shown that maintaining a higher RDI in the earlier treatment phase contributed to better prognosis.^{4,7} If a sufficiently high RDI is achieved until six cycles of standard regimen for DLBCL, additional chemotherapy may not have further prognostic benefit. Significantly more patients underwent HDT-ASCT in the six-cycle group than in the eight-cycle group, but fewer patients in the survivor group underwent HDT-ASCT. Patients requiring HDT-ASCT are usually limited to those with poor prognosis.²⁶ Notably, prognosis in the six-cycle group was not inferior to that in the eight-cycle group, even though a larger number of cases with poor prognosis in six cycles required HDT-ASCT.

Approximately one-third of DLBCL patients experience disease relapse or are refractory after initial standard therapy, but most elderly patients cannot benefit from HDT-ASCT because of ineligibility.²⁷ Therefore, in elderly patients who are not candidates for HDT-ASCT, treatment

with higher RDI in the initial treatment would be expected to improve prognosis. However, the present study showed that increasing the number of chemotherapy cycles from six to eight does not improve prognosis even in elderly patients. This is probably because patients requiring HDT-ASCT have an inherently poor prognosis, and R-standard chemotherapies alone do not improve the prognosis.²⁶ However, patients with high IPI are expected to represent a population in which eight treatment cycles would be superior, as six cycles might be inadequate to achieve better prognosis.²⁶ However, we note that six cycles of standard regimen are in no way inferior to eight cycles, independent of IPI, namely in any disease status. Patients with high IPI may require chemotherapy regimens more intensive than (R-)CHOP or THP-COP to improve prognosis.²⁸

The GOYA analysis, a prospective study, has recently also demonstrated non-inferiority of six cycles to eight cycles of R-CHOP.²⁹ The present study differs from GOYA analysis in the age group and the frailty of patients, a more real-world patient population than those in prospective trials, included in the analysis. The GOYA analysis, as well as other randomized controlled trials, has high internal validity. On the other hand, our trial using real-world data has high external validity, as it reflects the coming global aging society. Our results not only fill the evidence gap between the findings of GOYA analysis and clinical practice, but also provide further support for the findings of GOYA analysis. In the GOYA analysis, six and eight cycles have equal efficacy but eight cycles were associated with increased AEs.²⁹ In contrast, the occurrence of severe AEs between the two groups (six versus eight cycles) has no significant differences in the present study. The difference in the outcomes between our study and the GOYA analysis is possibly explained by the difference in the age group of the eligible patients between the two studies. In the present study, 71.0% of the patients were aged ≥ 65 years and there was no significant difference between the two groups (six versus eight cycles) in the G8 score in the population aged ≥ 65 years. We previously reported that the probability of AEs in the elderly DLBCL treated by the standard therapy is independent of tARDI and depends on the G8 score, the established and easy-to-use geriatric assessment tool mainly for the patients with solid cancers.¹⁹ Reflecting the real world, most of the patients in our study were elderly, thus the risk of AEs might be influenced by a host-depending factor, the G8 score, rather than the number of treatment cycles. Besides, the treatment-related mortality rate in our cohort (0.0%) was lower than in a previous report (6.0%).¹ This difference was attributed to the fact that most patients in Japan receive the initial treatment cycle as inpatients. Generous universal health insurance in Japan guaranteed sufficient supportive care including hospitalization.³⁰

The limitations of this study should be noted. Firstly, due to its retrospective nature, various unmeasured confounding factors may not have been addressed in propensity

score matching. Secondly, the results of this study cannot be considered to have been without selection bias. Patients treated with less than six cycles or with seven cycles of standard regimens were not included in this study. Thirdly, we cannot retrospectively identify which patients were planned to receive eight cycles but were terminated after six chemotherapy cycles because of AEs. Two options were available for physicians: administration of a defined number of cycles independent of response (six or eight cycles) or use of a response-based evaluation to determine the number of chemotherapy cycles. Whether the attending physician planned to administer six or eight cycles in advance was not taken into consideration. Intention-to-treat analysis thus could not be carried out. However, this flexibility in determining the number of treatment cycles allowed our data to reflect pragmatic clinical practice. Fourthly, we cannot exclude immortal time bias. We carried out sensitivity analyses to minimize or avoid immortal time bias and to address the concern of per protocol set analysis by the time-dependent Cox regression analysis and the interim analysis with sIL-2R, respectively.^{10,11} These sensitivity analyses, while important and supportive, cannot definitively address potential bias arising from the immortal bias and per protocol set analysis. However, when considering the effect of residual bias on the results, the primary finding from our study that prognosis was no better with eight cycles of (R-) CHOP or THP-COP-21 than with six cycles was consistent across the sensitivity analysis. The possibility that immortal time bias may have influenced the quantitative results remains, and caution should be exercised in extrapolating these findings to excluded population; however, the sensitivity analysis and the effect of residual bias suggest that immortal time bias and per protocol set analysis are unlikely to result in the direction of reversing the main results. Fifthly, the information about subtypes (GCB, ABC, or double-hit lymphoma) was not fully investigated in our study, thus further investigation is warranted to verify a potential influence on the outcome. Finally, these data were for a Japanese population, which is the most advanced 'super-aged' society in the world, but also reflects the real-world reality that many countries will be facing in the future.

In conclusion, at any RDI, no significant difference in OS was observed between six cycles of standard regimens as the initial treatment of DLBCL and eight cycles. Even after accounting for age and IPI modifications, eight cycles did not offer better prognosis than six cycles.

ACKNOWLEDGEMENTS

The authors acknowledge Dr Keiichi Kinoshita, Shin Imamura, and Yasukazu Kawai for their scientific advice. The authors also thank Ms Nami Fujita for her technical support.

FUNDING

None declared.

DISCLOSURE

TY has received research funds and honoraria from Stargen, Boehringer Ingelheim, Mundipharma, Solasia, Pfizer, and Chugai Pharmaceutical. All other authors have declared no conflicts of interest.

DATA SHARING

Data cannot be shared according to the protocol of the research participating institutions.

REFERENCES

1. Coiffier B, Lepage E, Briere J, et al. CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. *N Engl J Med*. 2002;346(4):235-242.
2. Pfreundschuh M, Schubert J, Ziepert M, et al. Six versus eight cycles of bi-weekly CHOP-14 with or without rituximab in elderly patients with aggressive CD20+ B-cell lymphomas: a randomised controlled trial (RICOVER-60). *Lancet Oncol*. 2008;9(2):105-116.
3. Wåsterlid T, Biczler JL, Brown PN, et al. Six cycles of R-CHOP-21 are not inferior to eight cycles for treatment of diffuse large B-cell lymphoma: a Nordic Lymphoma Group Population-based Study. *Ann Oncol*. 2018;29(8):1882-1883.
4. Kwak LW, Halpern J, Olshen RA, Horning SJ. Prognostic significance of actual dose intensity in diffuse large-cell lymphoma: results of a tree-structured survival analysis. *J Clin Oncol*. 1990;8(6):963-977.
5. Lee S, Fujita K, Negoro E, et al. Impact of relative dose intensity of standard regimens on survival in elderly patients aged 80 years and older with diffuse large B-cell lymphoma. *Haematologica*. 2020;105(8):e415-e418.
6. Bosly A, Bron D, Van Hoof A, et al. Achievement of optimal average relative dose intensity and correlation with survival in diffuse large B-cell lymphoma patients treated with CHOP. *Ann Hematol*. 2008;87(4):277-283.
7. Kanemasa Y, Shimoyama T, Sasaki Y, et al. The impacts of initial and relative dose intensity of R-CHOP on outcomes of elderly patients with diffuse large B-cell lymphoma. *Leuk Lymphoma*. 2017;58(3):736-739.
8. Swerdlow SH, Campo E, Harris NL, et al. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues*. 4th ed. Lyon, France: IARC Press; 2008.
9. Swerdlow SH, Campo E, Harris NL, et al. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. Revised*. 4th ed, Vol 2. Lyon, France: IARC Press; 2017.
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. Goto N, Tsurumi H, Goto H, et al. Serum soluble interleukin-2 receptor (sIL-2R) level is associated with the outcome of patients with diffuse large B cell lymphoma treated with R-CHOP regimens. *Ann Hematol*. 2012;91(5):705-714.
12. Dalia S, Chavez J, Little B, et al. Serum albumin retains independent prognostic significance in diffuse large B-cell lymphoma in the post-rituximab era. *Ann Hematol*. 2014;93(8):1305-1312.
13. International Non-Hodgkin's Lymphoma Prognostic Factors Project. A predictive model for aggressive non-Hodgkin's lymphoma. *N Engl J Med*. 1993;329(14):987-994.
14. Saygin C, Jia X, Hill B, et al. Impact of comorbidities on outcomes of elderly patients with diffuse large B-cell lymphoma. *Am J Hematol*. 2017;92(10):989-996.
15. Kanemasa Y, Shimoyama T, Sasaki Y, Hishima T, Omuro Y. Geriatric nutritional risk index as a prognostic factor in patients with diffuse large B cell lymphoma. *Ann Hematol*. 2018;97(6):999-1007.
16. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40(5):373-383.
17. Bouillanne O, Morineau G, Dupont C, et al. Geriatric Nutritional Risk Index: a new index for evaluating at-risk elderly medical patients. *Am J Clin Nutr*. 2005;82(4):777-783.
18. Kenis C, Decoster L, Van Puyvelde K, et al. Performance of two geriatric screening tools in older patients with cancer. *J Clin Oncol*. 2014;32(1):19-26.
19. Oiwa K, Fujita K, Lee S, et al. Utility of the Geriatric 8 for the prediction of therapy-related toxicity in older adults with diffuse large B-cell lymphoma. *Oncologist*. 2021;26(3):215-223.
20. Hara T, Yoshikawa T, Goto H, et al. R-THP-COP versus R-CHOP in patients younger than 70 years with untreated diffuse large B cell lymphoma: a randomized, open-label, noninferiority phase 3 trial. *Hematol Oncol*. 2018;36(4):638-644.
21. Common Terminology Criteria for Adverse Events (CTCAE). Version 4.0. May 28, 2009 (v4.03: June 14, 2010.). US Department of Health and Human Services; 2010.
22. Simon R, Makuch RW. A non-parametric graphical representation of the relationship between survival and the occurrence of an event: application to responder versus non-responder bias. *Stat Med*. 1984;3(1):35-44.
23. Shintani AK, Girard TD, Eden SK, Arbogast PG, Moons KG, Ely EW. Immortal time bias in critical care research: application of time-varying Cox regression for observational cohort studies. *Crit Care Med*. 2009;37(11):2939-2945.
24. Kanda Y. Investigation of the freely available easy-to-use software 'EZ' for medical statistics. *Bone Marrow Transplant*. 2013;48(3):452-458.
25. R Core Team. *Computing RFS. R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2016.
26. Stiff PJ, Unger JM, Cook JR, et al. Autologous transplantation as consolidation for aggressive non-Hodgkin's lymphoma. *N Engl J Med*. 2013;369(18):1681-1690.
27. Friedberg JW. Relapsed/refractory diffuse large B-cell lymphoma. *Hematology Am Soc Hematol Educ Program*. 2011;2011:498-505.
28. McMillan AK, Phillips EH, Kirkwood AA, et al. Favourable outcomes for high-risk diffuse large B-cell lymphoma (IPI 3-5) treated with front-line R-CODOX-M/R-IVAC chemotherapy: results of a phase 2 UK NCRI trial. *Ann Oncol*. 2020;31(9):1251-1259.
29. Sehn LH, Congiu AG, Culligan DJ, et al. No added benefit of eight versus six cycles of CHOP when combined with rituximab in previously untreated diffuse large B-cell lymphoma patients: results from the International Phase III GOYA study. *Blood*. 2018;132(suppl 1):783.
30. Tamiya N, Noguchi H, Nishi A, et al. Population ageing and wellbeing: lessons from Japan's long-term care insurance policy. *Lancet*. 2011;378(9797):1183-1192.