BMJ Open

Risk and Outcome of Acute Myeloid Leukemia Among Survivors of Primary Diffuse Large B-Cell Lymphoma: a retrospective observational study based on SEER database

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2022-061699
Article Type:	Original research
Date Submitted by the Author:	09-Feb-2022
Complete List of Authors:	Du, Yu; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Wang, Ying; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Li, Qinlu; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Chang, Xiaona; Wuhan Union Hospital Zhang, Heng; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Xiao, Min; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology xing, shugang; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology,
Keywords:	EPIDEMIOLOGY, Leukaemia < HAEMATOLOGY, Lymphoma < HAEMATOLOGY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Risk and Outcome of Acute Myeloid Leukemia Among Survivors of Primary Diffuse Large B-Cell Lymphoma: a retrospective observational study based on SEER database

Yu Du[1], Ying Wang[2], Qinlu Li[2], Xiaona Chang[3], Heng Zhang[2], Min Xiao[2], Shugang Xing[2]

1 Department of Radiology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China.
2 Department of Hematology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China.
3 Department of Pathology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China.

Corresponding Author: Shugang Xing, Department of Hematology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, 1095 JieFang Avenue, Wuhan 430030, China (xingsg@tjh.tjmu.edu.cn).

Abstract

Objectives Survivors of diffuse large B-cell lymphoma (DLBCL) are at an increased risk of developing second primary malignancies (SPMs). However, the risk of secondary acute myeloid leukemia (sAML) has not been previously described in detail, and the outcomes of patients with sAML are also undiscovered compared with their de novo counterparts (dnAML).

Design This study is a retrospective database study.

Setting and participants A total of 70280 patients with primary DLBCL, diagnosed between 2000 and 2016, were identified from the Surveillance, Epidemiology, and End Results (SEER) database. Another cohort with dnAML matching with sAML was also obtained from SEER database.

Results The standardized incidence ratio was 6.23 (95% confidence interval [CI], 5.50–7.03) for sAML among survivors of DLBCL. The estimated cumulative incidence of sAML was 0.61% 15 years after the diagnosis of DLBCL. Patients aged 60–74 years were more likely to have sAML than those <60 years (sHR=1.417; 95% CI, 1.087–1.850), whereas patients aged ≥75 years were less likely to have sAML (sHR=0.648; 95% CI, 0.452–0.930). Patients with advanced-stage DLBCL were more prone to sAML than those with early-stage disease (sHR=1.307; 95% CI, 1.012–1.690). There was a significant difference of survival between patients with dnAML and those with sAML (hazard ratio=1.25; 95% CI, 1.01–1.53).

Conclusions The risk of developing sAML after DLBCL is substantial. Patients aged 60–74 years and with advanced-stage are more prone to sAML. And, compared to their dnAML counterparts, patients with sAML have a worse prognosis.

Keywords diffuse large B-cell lymphoma; secondary acute myeloid leukemia; standardized incidence ratio; competing risk model; hazard factor; overall survival

Strengths and limitations of this study

- Surveillance, Epidemiology, and End Results (SEER) database is a large database of US patients, but the detailed information regarding disease treatments is not mentioned.
- Competing risk model was performed to eliminate the effect of death, which would lead to a bias for the incidence of sAML.
- Case matching analysis was performed to eliminate the effect of confounding variables between sAML group and dnAML group.

Introduction

Diffuse large B-cell lymphoma (DLBCL) is the most common and aggressive type of lymphoma.[1] The combination of rituximab with CHOP (cyclophosphamide, doxorubicin, vincristine and prednisolone) chemotherapy has improved the overall survival (OS) of patients with DLBCL by at least 20%.[2] However, with the increasing survival of patients after DLBCL, the risk of second primary malignancies (SPMs) has also increased, and their management has become an emerging challenge. Currently, SPMs are an important cause of death among survivors of DLBCL.[3,4]

One of the main secondary malignancies following DLBCL is acute myeloid leukemia (AML). For years, more cases of AML have been reported in survivors DLBCL than in the general population.[5,6] Although the underlying factors and biological mechanisms of AML following DLBCL need to be better clarified, the factors about treatment, including the use of rituximab, have been thought to be the main cause of the increased risk.[7-9] The management of patients with secondary AML (sAML) may be challenging because of cumulative toxicity from the treatment of primary DLBCL. Previous exposure to treatment-related factors, including radiotherapy and systemic chemotherapy, has limited the treatment options for secondary neoplasms and further alters their outcomes.[10-13] Hence, we asked if the outcome of secondary AML was poorer than that of de novo AML (dnAML). Meanwhile, considering the difficulty in the management of patients with sAML, a search for the predictive factors for the occurrence of sAML will be meaningful. As far as we know, information available regarding the incidence and prognosis of sAML following DLBCL is limited.

Here, we used sequential cancer data available from the large and high-quality, population-based National Cancer Institute Surveillance, Epidemiology, and End Results (SEER) program to describe the pattern of incidence, investigate the predictive factors for the occurrence of sAML, and compare the outcomes of patients with sAML with their de novo counterparts.

Materials and Methods

Data Source and Sample

The SEER program's research data for 17 registries (excluding Alaska) were used to assess the incidence and explore the hazard factors of sAML in survivors of primary DLBCL diagnosed between 2000 and 2016. DLBCL cases were identified according to the Lymphoma Subtype Recode/WHO 2008, which is updated for hematopoietic conditions and coded based on ICD-O-3 and the WHO Classification of Tumors of Hematopoietic and Lymphoid Tissues (2008).[14] We excluded cases which were coded as autopsy or death-certificate-only, where DLBCL was not the first primary cancer, and those with unknown age or race. To exclude patients with synchronous DLBCL and AML, cases diagnosed with AML within the first 2 months of being diagnosed with

DLBCL were not included in this study, as well as those with <2 months of follow-up. The stage at diagnosis of DLBCL was classified into two: stage Ⅰ and Ⅱ disease as early stage, and stage Ⅲ and Ⅳ disease as advanced stage.

Statistical Analysis

Continuous numerical variables were described by median and range, and comparison between the two groups was done with the Wilcoxon rank-sum nonparametric test. Differences in proportions across the groups were compared with the chi-square test. The calculation of the standardized incidence ratio (SIR) and 95% confidence interval (CI) for sAML in patients with DLBCL was performed using SEER*stat software (version 8.3.6, NCI, NIH, Bethesda, MD, USA). And, the calculation of total person-years and person-years in each subgroup of DLBCL survivors was also performed by SEER*stat.

The analyses of cumulative incidence of sAML were completed using competing risk models, in which death from any cause was considered the sole competing risk. The differences in cumulative incidence among the groups were compared using Gray's test. Furthermore, to explore the risk factors for sAML, a regression analysis using the semiparametric proportional hazards model proposed by Fine and Gray was performed.[15] Using these models, the semiparametric hazard ratios (sHRs) and their 95% CIs for risk factors were estimated.

Furthermore, in order to explore whether there was a difference in survival outcome between patients with sAML and patients with dnAML, we first listed the detailed characteristics of all patients with sAML. We then obtained cast lists of dnAML from the SEER database using the same histological subtype as for cases of sAML, but the AML was the first malignancy for a given individual. Then, based on age (±2 years), calendar year of diagnosis (±2 years), sex, and race, we matched sAML with dnAML patients at a 1:1 ratio. Case matching was completely random and the variables (survival status and cause of death) that might affect the matching result were with no awareness. Because the SEER database does not have detailed information about the treatment, matching for therapy was impossible. We used a shared-frailty Cox model to interpret the 1:1 matched design. Meanwhile, the factors, age, sex, race, and number of years of diagnosis were adjusted for the model. For AML with previous DLBCL versus dnAML, the hazard ratio (HR) and its 95% CI was calculated.

R software (version 3.6.3) with "cmprsk" and "survival" packages and STATA (version 14.0; Stata Corporation, College Station, TX, USA) were used to perform these analyses. In this study, we treated a two-sided P value < 0.05 to be a statistically significant difference.

Patient and public involvement

No patients or public were involved in this study.

Results

In this study, we identified a total of 70280 primary DLBCL patients, and the median follow-up is 90 months (range, 2–203 months), contributing to a total follow-up of 334,516 person-years. By the end of the follow-up, 264 of these cases having a diagnosis of sAML. The median interval between the diagnosis of DLBCL and sAML was 44 months (range, 3–178 months). The characteristics of the entire cohort of patients with DLBCL who have or have not developed sAML are shown in Table 1.

Table 1. Characteristics of two-month survivors of DLBCL reported to the SEER Program (2000-2016)

Characteristic	All n=70280	No sAML n=700016	sAML n=264	*P*
Follow-up (range), mo	90 (2-203)	90 (2-203)	157 (3-198)	
Age (range), yrs	64 (0-106)	64 (0-106)	63.5 (12-88)	0.197
Age group, yrs				< 0.001
< 60	28289 (40.3%)	28186 (40.3%)	103 (39.0%)	
60-74	23796 (33.9%)	23677 (33.8%)	119 (45.1%)	
75+	18195 (25.9%)	18153 (25.9%)	42 (15.9%)	
Sex				0.165
Male	38409 (54.7%)	38253 (54.6%)	156 (59.1%)	
Female	31871 (45.3%)	31763 (45.4%)	108 (40.9%)	
Race				0.490
Black	5499 (7.8%)	5483 (7.8%)	16 (6.1%)	
White	58661 (83.5%)	58434 (83.5%)	227 (86.0%)	
Other	6120 (8.7%)	6099 (8.7%)	21 (8.0%)	
Primary site				0.015
Nodal	46241 (65.8%)	46048 (65.8%)	193 (73.1%)	
Extranodal	24039 (34.2%)	23968 (34.2%)	71 (26.9%)	
Ann Arbor stage				0.001
Stage Ⅰ	18535 (26.4%)	18470 (26.4%)	65 (24.6%)	
Stage Ⅱ	13717 (19.5%)	13672 (19.5%)	45 (17.0%)	
Stage Ⅲ	10726 (15.3%)	10672 (15.2%)	54 (20.5%)	
Stage Ⅳ	20026 (28.5%)	19937 (28.5%)	89 (33.7%)	
Unknown	7276 (10.4%)	7265 (10.4%)	11 (4.2%)	
Years of diagnosis				< 0.001
2000-2005	23047 (32.8%)	22933 (32.8%)	114 (43.2%)	
2006-2011	25196 (35.9%)	25084 (35.8%)	112 (42.4%)	
2012-2016	22037 (31.4%)	21999 (31.4%)	38 (14.4%)	

Abbreviations: DLBCL, diffuse large B-cell lymphoma; sAML, secondary acute myeloid leukemia

The SIR for sAML overall was 6.23 (95% CI, 5.50–7.03), indicating an elevated incidence compared with that for the general population of the USA. The forest plot for the SIRs is shown in Figure 1. The SIR was 13.46 (95% CI, 10.99–16.33) in patients aged <60 years, 6.17 (95% CI, 5.11–7.39) in patients aged 60–74 years, and 2.72 (95% CI, 1.96–3.68) in patients aged ≥75 years; thus, it decreased with increasing age (*P* for trend < 0.001). The nodal DLBCL (SIR=7.08; 95% CI, 6.12–8.15) had a higher SIR for sAML than extranodal DLBCL (SIR=4.70; 95% CI, 3.67–5.93), *P*=0.003. As for the Ann Arbor Stage of DLBCL, the SIR was less for the early-stage (SIR=4.86; 95% CI, 3.99–5.86) as compared to that for advanced-stage disease (SIR=8.38; 95% CI, 7.07–9.88) (*P*<0.001). Patients with a latency of 24–59 months had a higher SIR (8.31; 95%CI, 6.82–10.04) than those with other latencies. For the groups of sex, race, and years of diagnosis, no heterogeneity or trend for SIRs was observed.

Further, when competing causes of deaths were considered, the cumulative incidence of sAML was 0.30% (95% CI, 0.26%–0.35%), 0.53% (95% CI, 0.46%–0.60%), and 0.61% (95% CI, 0.53%–0.70%) at 5, 10, and 15 years after the diagnosis of DLBCL, respectively. Moreover, we found that the cumulative incidence of sAML was closely related to the patients' age at the diagnosis of DLBCL (*P*<0.001), the primary site (*P*=0.010), and the Ann Arbor stage of DLBCL (*P*=0.007). The cumulative incidence at 10 years after DLBCL diagnosis was 0.51% (95% CI, 0.41%–0.62%) in patients aged <60 years, 0.74% (95% CI, 0.61%–0.89%) in patients aged 60–74 years, and 0.29% (95% CI, 0.21%–0.39%) in patients aged ≥75 years. For extranodal DLBCL, the cumulative incidence in patients at 10 years was 0.40% (95% CI, 0.31%–0.51%); it was 0.59% (95% CI, 0.51%–0.69%) for DLBCL occurring in the lymph node. As regards the Ann Arbor Stage of DLBCL, the cumulative incidence in patients at 10 years was 0.43% (95% CI,

0.35%–0.52%) and 0.66% (95% CI, 0.55%–0.79%) for the early and advanced stages, respectively (Figure 2).

Furthermore, according to the semiparametric proportional hazards model, we investigated the risk factors for sAML occurrence. The results are presented in Table 2. Univariate analyses showed that patients' age, primary site, and Ann Arbor stage of DLBCL were statistically significant risk factors (P<0.05). These three variables were selected for the final multivariate analysis, which showed that the patients' age at diagnosis and the Ann Arbor stage of DLBCL were independent predictors of the occurrence of sAML. Patients aged 60–74 years were more easily to have sAML (sHR=1.417; 95% CI, 1.087–1.850; P=0.010) than those aged <60 years. However, patients aged ≥75 years were less likely to have sAML than patients aged <60 years (sHR=0.648; 95% CI, 0.452–0.930; P=0.018). Patients with advanced-stage DLBCL were more prone to sAML than those with early-stage disease (sHR=1.307; 95% CI, 1.012–1.690; P=0.040).

Table 2. Univariate and multivariate analyses for predictive factors of sAML.

Factors	Univariate			Multivariate		
	sHR	95% CI	*P*	sHR	95% CI	*P*
Age, yrs						
< 60	ref.		0.009	ref.		
60-74	1.421	1.092-1.850		1.417	1.087-1.850	0.010
75+	0.635	0.443-0.908	0.013	0.648	0.452-0.930	0.018
Sex						
Male	ref.					
Female	0.827	0.647-1.060	0.130			
Race						
White	ref.					
Black	0.764	0.460-1.270	0.300			
Other	0.947	0.605-1.480	0.810			
Primary site						
Nodal	ref.			ref.		
Extranodal	0.703	0.536-0.923	0.011	0.770	0.583-1.020	0.065
Ann Arbor stage						
Early stage	ref.			ref.		
Advanced stage	1.423	1.110-1.830	0.005	1.307	1.012-1.690	0.040
Unknown	0.808	0.434-1.500	0.500	0.807	0.431-1.510	0.500

Abbreviations: sAML, secondary acute myeloid leukemia; sHR, subdistribution hazard ratios; ref, reference

The Table 3 listed the characteristics of patients with sAML and their dnAML counterparts. The median survival time for patients with sAML and dnAML was 7 months (95% CI, 6–9 months) and 13 months (95% CI, 10–17 months), respectively (Figure 3). The Cox model showed that patients with sAML had a higher risk of death and a shorter OS than their dnAML counterparts (HR=1.25; 95% CI, 1.01–1.53; *P*=0.038). Of all the causes of death, AML is the most common in patients with both sAML and dnAML. However, we found that death from DLBCL was still a main component of overall mortality for patients who subsequently developed sAML. (Table 4).

Table 3. The baseline characteristics of patients with sAML and the matched cases with dnAML.

Characteristics	sAML n=262	dnAML n=262
Age (range), yrs	68.0 (15.0-95.0)	67.5 (15.0-93.0)
Age group, yrs		
< 60	67 (25.6%)	71 (27.1%)
60-74	121 (46.2%)	120 (45.8%)
75+	74 (28.2%)	71 (27.1%)
Sex		
Male	155 (59.2%)	155 (59.2%)
Female	107 (40.8%)	107 (40.8%)
Race		
Black	16 (6.1%)	16 (6.1%)
White	225 (85.9%)	225 (85.9%)
Other	21 (8.0%)	21 (8.0%)
Year of diagnosis		
Median (range)	2011 (2001-2016)	2010 (2000-2016)

Abbreviations: sAML, secondary acute myeloid leukemia; dnAML, de novo AML

Table 4. Causes of death in patients with sAML and matched dnAML

Causes of Death	sAML (n)	dnAML (n)
No. of deaths	218	210
AML	105	144
DLBCL	66	0
Other cancer	18	21
Other Hematopoietic and lymphoid tumors	12	17
Solid tumor	6	4
Cardiovascular and cerebrovascular	5	6
Infection	4	6
Other	20	28
NA	0	5

Abbreviations: sAML, secondary acute myeloid leukemia; dnAML, de novo AML; AML, acute myeloid leukemia; NHL, non-Hodgkin lymphoma; NA, not available

Discussion

As far as we know, this is the largest population-based study of sAML in patients with DLBCL. In this study, we observed an increased incidence of sAML among survivors of DLBCL and demonstrated substantial heterogeneity in the occurrence of sAML by age at diagnosis, primary site and Ann Arbor stage of DLBCL. Specifically, we identified that the age at diagnosis and stage of DLBCL were independent risk factors for sAML. We also observed that sAML had a shorter OS than dnAML, and that death from DLBCL was a main component of overall mortality for patients who subsequently developed sAML.

In a population-based study, the SIR of sAML was 4.29 for patients with DLBCL, indicating a higher incidence of sAML in patients with DLBCL than that in the general population,[16] which is consistent with our results. Another large study combined data from 25,089 patients with DLBCL from California and reported 75 cases of sAML.[7] This was 4.39-(pre-rituximab) or 8.70-(post-rituximab) times the number of expected cases from the general population, indicating an increased risk, which was similar to that reported herein.

In this study, we confirmed that the SIR of sAML decreased with an increase in age at diagnosis. However, when competing causes of death were considered, patients aged 60–74 years had the highest cumulative probability of sAML at 10 years of follow-up. This result is consistent with that reported in the papers, which show that sAML tends to occur more commonly in older patients.[17–19] However, this study also showed that

patients aged ≥75 years had a lower cumulative incidence than younger patients. Since high-dose chemoradiotherapy has been associated with an elevated risk of sAML,[20] we speculate that the older patients may not be able to tolerate high-dose chemoradiotherapy and may not be suitable candidates for intensive treatment for primary DLBCL. Nevertheless, we could not obtain detailed data on the treatment of DLBCL from the SEER database for further research. The semiparametric proportional hazards model also showed that age 60–74 years was an independent risk factor for sAML, and that age ≥75 years was less easily to have sAML compared with age <60 years in this study.

Together with age, our study identified two other potential risk factors for sAML among survivors of patients with DLBCL: primary site and the Ann Arbor stage of DLBCL.

The link of AML risk with the stage at diagnosis of DLBCL has been not well clarified for DLBCL. A large population-based study indicated that patients with advanced-stage DLBCL were more likely to develop hematological SPMs, and that the most common histology of hematological SPM was AML.[6] In this study, we also found that patients with advanced-stage DLBCL had a higher SIR than those with early-stage DLBCL (P<0.001). Considering the competing causes of death, we found that the cumulative probability of sAML for patients with advanced-stage DLBCL was higher than that for patients with early-stage at 10 years of follow-up. Furthermore, the advanced-stage DLBCL was identified as an independent risk factor for sAML in our study.

We also indicated that early-stage DLBCL may show a unique biology compared with advanced-stage DLBCL. In an experimental research with the analysis of gene expression profiling, a number of genes were differentially expressed in patients with early-stage DLBCL compared to those with advanced-stage DLBCL.[21] Another study suggested that increased late relapses in the early-stage DLBCL compared with advanced-stage DLBCL may be caused by biological differences.[22] However, a report of patients with advanced-stage DLBCL also recognized the risk of late relapse.[23] Thus, further study of the biological differences between early- and advanced-stage DLBCL is needed.

Our study found that patients with primary sites in the lymph nodes had a higher SIR and cumulative probability than those with extranodal disease. However, according to the semiparametric proportional hazards model, multivariate analysis showed that the primary sites of DLBCL were not independent risk factors for the progress of sAML. As reported in the literature, patients with early-stage DLBCL are more likely to have extranodal disease.[24] This study also showed similar results (data not shown). Given this finding, it is possible that early-stage DLBCL, which is mainly located in extranodal sites, lowers the SIR or cumulative probability.

In the present study, we compared the survival outcomes of patients with sAML and

their dnAML counterparts. The results show that the prognosis is worse for patients diagnosed with sAML after surviving DLBCL than those diagnosed with dnAML in matched cases. Previously studies have indicated that the prior therapy of DLBCL shows a detrimental effect, which has been verified in patients who have developed malignant mesotheliomas, bladder cancer, and kidney cancer.[25–27] The successfully treat of second cancer in patients who survive DLBCL has been affected by many factors, such as limitations on the dose and site of radiotherapy, a poor tolerance to chemotherapy, and impaired physiologic reserve. Another intriguing factor may result from the intrinsically worse biology of sAML, which require more in-depth research.

The primary limitation of this study is that we cannot obtain the detailed information regarding disease treatments. Therefore, it is impossible to establish a correlation between DLBCL treatment and the development of sAML. In addition, the therapeutic modalities that could be used to treat patients with sAML are also not mentioned in the database, which limits the exploration of prognosis. However, there are several novel findings shown in this study. These findings may be helpful in future prospective trials for patients with DLBCL.

In conclusion, the current findings suggest that the incidence of AML increases significantly among survivors of DLBCL. Furthermore, we showed that age and Ann Arbor stage of DLBCL at diagnosis are independent risk factors for sAML. We also found that patients with sAML had shorter OS than their dnAML counterparts. These findings will be beneficial for the management of patients with newly diagnosed DLBCL.

Acknowledgment

We thank Editage (www.editage.cn) for their help in English language editing.

Funding

No funding for this study.

Declaration of Interest

The authors declare no potential conflicts of interest.

Patient consent for publication

Not applicable.

Ethics approval

This study used the SEER research database, which is approved by the NIH Ethics Program.

Authors Contributions

Yu Du: Conceptualization and design, collection and assembly of the data, formal analysis and interpretation, visualization, writing-original draft, and writing-review and editing. **Ying Wang:** Collection and assembly of the data, data curation, visualization, writing-original draft, and writing-review and editing. **Qinlu Li:** Data curation and analysis, writing-review and editing. **Xiaona Chang:** Methodology, supervision, and writing-review. **Heng Zhang:** Supervision and writing-review. **Min Xiao:** Conceptualization, methodology, validation, data curation and writing-review and editing. **Shugang Xing:** Conceptualization, methodology, validation, formal analysis, data curation, supervision, writing-original draft, writing-review and editing.

Data availability statement

Based on website https://www.cancer.gov/policies/accessibility, the National Cancer Institute (NCI) provides access to all individuals seeking information on http://www.cancer.gov. Besides, on the website https://www.cancer.gov/policies/copyright-reuse, we are also informed that most of the information on NCI website is in the public domain and is not subject to copyright restrictions. No special permission is required to use or reproduce public domain material.

References

1. Swerdlow SH, Campo E, Harris NL, et al, eds. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues*. 4th ed. IARC Press; 2008
2. 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;346:235–242. doi: 10.1056/NEJMoa011795
3. de Moor JS, Mariotto AB, Parry C, et al. Cancer survivors in the United States: Prevalence across the survivorship trajectory and implications for care. *Cancer Epidemiol Biomarkers Prev.* 2013;22:561-570. doi: 10.1158/1055-9965.EPI-12-1356
4. Chinen Y, Tanba K, Takagi R, et al. Second primary malignancy after rituximab-containing immunochemotherapy for diffuse large B cell lymphoma. *Leuk Lymph*. 2020; 61(14):3378-3386. doi: 10.1080/10428194.2020.1811862
5. Gangavarapu LS, Surapaneni R. Diffuse large B-cell lymphoma survivorship: Risk of second malignancies in long term survivors. *Blood.* 2015;126(23):3921. doi: 10.1182/blood.V126.23.3921.3921
6. Major A, Smith DE, Ghosh D, et al. Risk and subtypes of secondary primary malignancies in diffuse large B-cell lymphoma survivors change over time based on stage at diagnosis. *Cancer*. 2020;126(1):189-201. doi: 10.1002/cncr.32513
7. Tao L, Clarke CA, Rosenberg AS, et al. Subsequent primary malignancies after diffuse large B-cell lymphoma in the modern treatment era. *Br J Haematol*. 2017;178(1):72-80. doi: 10.1111/bjh.14638
8. Fleury I, Chevret S, Pfreundschuh M, et al. Rituximab and risk of second primary malignancies in patients with non-Hodgkin lymphoma: a systematic review and meta-analysis. *Ann Oncol*. 2016;27(3):390-397. doi: 10.1093/annonc/mdv616
9. Brown JR, Yeckes H, Friedberg JW, et al. Increasing incidence of late second malignancies after conditioning with cyclophosphamide and total-body irradiation and autologous bone marrow transplantation for non-Hodgkin's lymphoma. *J Clin Oncol*. 2005;23(10):2208-2214. doi: 10.1200/JCO.2005.05.158
10. Milano MT, Li H, Constine LS, et al. Survival after second primary lung cancer: A population-based study of 187 Hodgkin lymphoma patients. *Cancer*. 2011;117(24):5538-5547. doi: 10.1002/cncr.26257
11. Milano MT, Li H, Gail MH, et al. Long-term survival among patients with Hodgkin's lymphoma who developed breast cancer: a population-based study. *J Clin Oncol.* 2010;28(34):5088-5096. doi: 10.1200/JCO.2010.29.5683
12. Hemminki K, Lenner P, Sundquist J, et al. Risk of subsequent solid tumors after non-Hodgkin's lymphoma: effect of diagnostic age and time since diagnosis. *J Clin Oncol*. 2008;26(11):1850-1857. doi: 10.1200/JCO.2007.14.6068.
13. Herr MM, Schonfeld SJ, Dores GM, et al. Risk for malignancies of infectious etiology among adult survivors of specific non-Hodgkin lymphoma subtypes. *Blood Adv*. 2019;3(13):1961-1969. doi: 10.1182/bloodadvances.2019030924
14. Sabattini E , Bacci F , Sagramoso C, et al. WHO classification of tumors of haematopoietic and lymphoid tissues in 2008: An overview.

Pathologica.2008;102(3):83-87.
15. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc*. 1999;94(446):496-509.
16. Jiang S, Zhen H, Jiang H. Second primary malignancy in diffuse large B-cell lymphoma patients: A SEER database analysis. *Curr Probl Cancer*. 2020;44(1):100502. doi: 10.1016/j.currproblcancer.2019.100502
17. Szotkowski T, Rohon P, Zapletalova L, et al. Secondary acute myeloid leukemia - a single center experience. *Neoplasma*. 2010;57(2):170-178. doi: 10.4149/neo_2010_02_170
18. Ossenkoppele G, Montesinos P. Challenges in the diagnosis and treatment of secondary acute myeloid leukemia. *Crit Rev Oncol Hematol*. 2019;138:6-13. doi: 10.1016/j.critrevonc.2019.03.003
19. Yang D, Fu X, Zhang X, et al. Therapy-related acute myeloid leukemia in patients with lymphoma: A report of four cases and review of the literature. *Oncol Lett*. 2015;10(5):3261-3265. doi: 10.3892/ol.2015.3703
20. Metayer C, Curtis RE, Vose J, et al. Myelodysplastic syndrome and acute myeloid leukemia after autotransplantation for lymphoma: a multicenter case-control study. *Blood.* 2003;101(5):2015-2023. doi: 10.1182/blood-2002-04-1261
21. Castillo JJ, Winer ES, Olszewski AJ. 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*. 2014;89(3):310-314. doi: 10.1002/ajh.23638
22. Roberts RA, Rimsza LM, Staudt L, et al. Gene expression differences between low and high stage diffuse large B cell lymphoma (DLBCL). *Blood*. 2006;108(11):809. doi: 10.1182/blood.V108.11.809.809
23. Stephens DM, Li H, LeBlanc ML, et al. Continued Risk of Relapse Independent of Treatment Modality in Limited-Stage Diffuse Large B-Cell Lymphoma: Final and Long-Term Analysis of Southwest Oncology Group Study S8736. *J Clin Oncol*. 2016;34(25):2997-3004. doi: 10.1200/JCO.2015.65.4582
24. Coiffier B, Thieblemont C, Van Den Neste E, et al. Long-term outcome of patients in the LNH-98.5 trial, the first randomized study comparing rituximab-CHOP to standard CHOP chemotherapy in DLBCL patients: a study by the Groupe d'Etudes des Lymphomes de l'Adulte. *Blood*. 2010;116(12):2040-2045. doi: 10.1182/blood-2010-03-276246
25. Chirieac LR, Barletta JA, Yeap BY, et al. Clinicopathologic characteristics of malignant mesotheliomas arising in patients with a history of radiation for Hodgkin and non-Hodgkin lymphoma. *J Clin Oncol*. 2013;31(36):4544-4549. doi: 10.1200/JCO.2013.49.9616
26. Travis LB, Curtis RE, Glimelius B, et al. Bladder and kidney cancer following cyclophosphamide therapy for non-Hodgkin's lymphoma. J Natl Cancer Inst. 1995;87(7):524-530. doi: 10.1093/jnci/87.7.524
27. André M, Mounier N, Leleu X, et al. Second cancers and late toxicities after treatment of aggressive non-Hodgkin lymphoma with the ACVBP regimen: a GELA cohort study on 2837 patients. Blood. 2004;103(4):1222-1228. doi: 10.1182/blood-

2003-04-1124

Figure Legends:

Figure 1. The standardized incidence ratio (SIR) forest plot for patients with secondary acute myeloid leukemia among survivors of diffuse large B-cell lymphoma. 95% CI, 95% confidence interval; ref, reference.

Figure 2. Cumulative incidence of secondary acute myeloid leukemia among survivors of diffuse large B-cell lymphoma. AML, acute myeloid leukemia.

Figure 3. The comparative outcome between survivors of diffuse large B-cell lymphoma who developed secondary acute myeloid leukemia (sAML) and matching patients with de novo acute myeloid leukemia (dnAML).

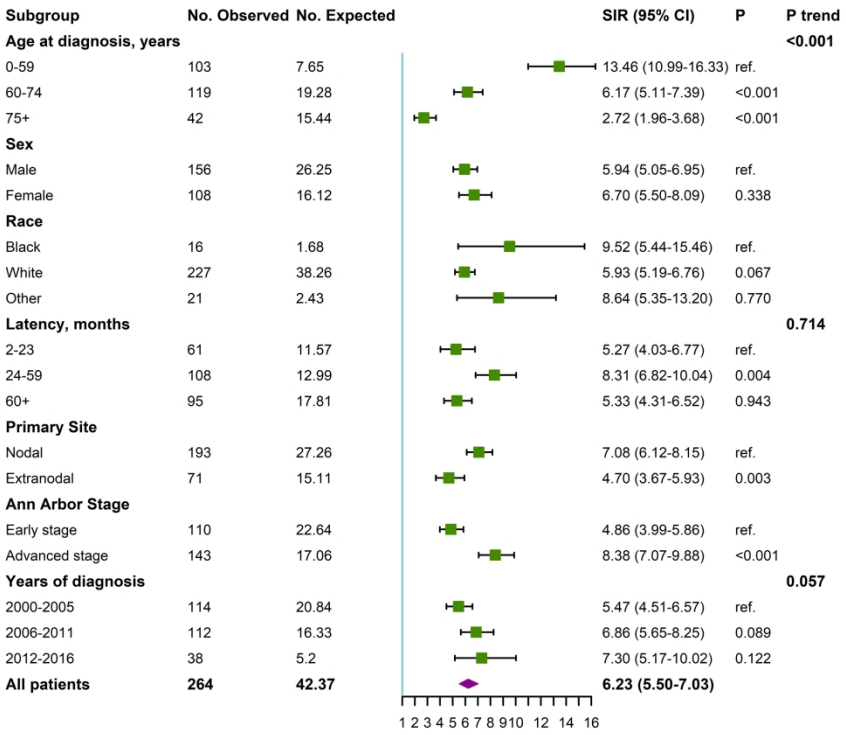

Figure 1. The standardized incidence ratio (SIR) forest plot for patients with secondary acute myeloid leukemia among survivors of diffuse large B-cell lymphoma. 95% CI, 95% confidence interval; ref, reference.

254x217mm (300 x 300 DPI)

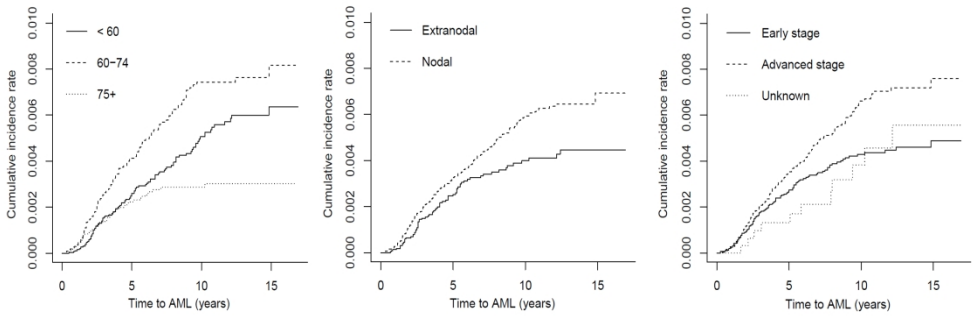

Figure 2. Cumulative incidence of secondary acute myeloid leukemia among survivors of diffuse large B-cell lymphoma. AML, acute myeloid leukemia.

621x241mm (150 x 150 DPI)

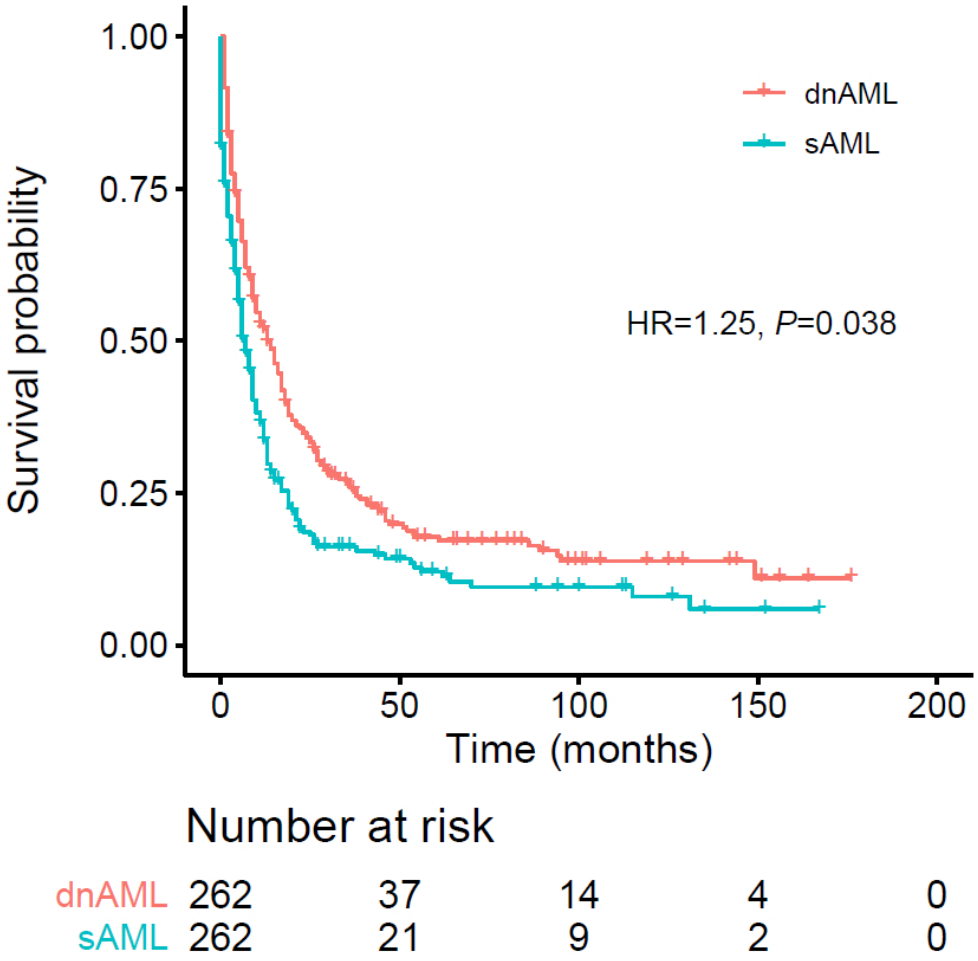

Figure 3. The comparative outcome between survivors of diffuse large B-cell lymphoma who developed secondary acute myeloid leukemia (sAML) and matching patients with de novo acute myeloid leukemia (dnAML).

STROBE Statement—Checklist of items that should be included in reports of ***case-control studies***

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1-2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	1-2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3
Objectives	3	State specific objectives, including any prespecified hypotheses	3
Methods			
Study design	4	Present key elements of study design early in the paper	3
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	3
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls	3
		(*b*) For matched studies, give matching criteria and the number of controls per case	4
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	4
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	4
Bias	9	Describe any efforts to address potential sources of bias	4
Study size	10	Explain how the study size was arrived at	4
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	4
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	4
		(*b*) Describe any methods used to examine subgroups and interactions	4
		(*c*) Explain how missing data were addressed	4
		(*d*) If applicable, explain how matching of cases and controls was addressed	4
		(*e*) Describe any sensitivity analyses	4
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	5
		(b) Give reasons for non-participation at each stage	5
		(c) Consider use of a flow diagram	5
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	5
		(b) Indicate number of participants with missing data for each variable of interest	5
Outcome data	15*	Report numbers in each exposure category, or summary measures of exposure	6

Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	6
		(*b*) Report category boundaries when continuous variables were categorized	6
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	6
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	7
Discussion			
Key results	18	Summarise key results with reference to study objectives	9
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	10
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	9
Generalisability	21	Discuss the generalisability (external validity) of the study results	9
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	11

*Give information separately for cases and controls.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

BMJ Open

Risk and Outcome of Acute Myeloid Leukemia Among Survivors of Primary Diffuse Large B-Cell Lymphoma: a retrospective observational study based on SEER database

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2022-061699.R1
Article Type:	Original research
Date Submitted by the Author:	06-Jul-2022
Complete List of Authors:	Du, Yu; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Wang, Ying; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Li, Qinlu; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Chang, Xiaona; Wuhan Union Hospital Zhang, Heng; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology Xiao, Min; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology xing, shugang; Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology,
Primary Subject Heading:	Haematology (incl blood transfusion)
Secondary Subject Heading:	Epidemiology, Oncology
Keywords:	EPIDEMIOLOGY, Leukaemia < HAEMATOLOGY, Lymphoma < HAEMATOLOGY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Risk and Outcome of Acute Myeloid Leukemia Among Survivors of Primary Diffuse Large B-Cell Lymphoma: a retrospective observational study based on SEER database

Yu Du[1], Ying Wang[2], Qinlu Li[2], Xiaona Chang[3], Heng Zhang[2], Min Xiao[2], Shugang Xing[2]

1 Department of Radiology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China.
2 Department of Hematology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China.
3 Department of Pathology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China.

Corresponding Author: Shugang Xing, Department of Hematology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, 1095 JieFang Avenue, Wuhan 430030, China (xingsg@tjh.tjmu.edu.cn).

Abstract

Objectives Survivors of diffuse large B-cell lymphoma (DLBCL) are at an increased risk of developing second primary malignancies (SPMs). However, the risk of secondary acute myeloid leukemia (sAML) has not been previously described in detail, and the outcomes of patients with sAML are also undiscovered compared with their de novo counterparts (dnAML).

Design This study is a retrospective database study.

Setting and participants A total of 70280 patients with primary DLBCL, diagnosed between 2000 and 2016, were identified from the Surveillance, Epidemiology, and End Results (SEER) database. Another cohort with dnAML matching with sAML was also obtained from SEER database.

Results The standardized incidence ratio was 6.23 (95% confidence interval [CI], 5.50–7.03) for sAML among survivors of DLBCL. The estimated cumulative incidence of sAML was 0.61% 15 years after the diagnosis of DLBCL. Patients aged 60–74 years were more likely to have sAML than those <60 years (sHR=1.417; 95% CI, 1.087–1.850), whereas patients aged ≥75 years were less likely to have sAML (sHR=0.648; 95% CI, 0.452–0.930). Patients with advanced-stage DLBCL were more prone to sAML than those with early-stage disease (sHR=1.307; 95% CI, 1.012–1.690). There was a significant difference of survival between patients with dnAML and those with sAML (hazard ratio=1.25; 95% CI, 1.01–1.53).

Conclusions The risk of developing sAML after DLBCL is substantial. Patients aged 60–74 years and with advanced-stage are more prone to sAML. And, compared to their dnAML counterparts, patients with sAML have a worse prognosis.

Keywords diffuse large B-cell lymphoma; secondary acute myeloid leukemia; standardized incidence ratio; competing risk model; hazard factor; overall survival

Strengths and limitations of this study

- Surveillance, Epidemiology, and End Results (SEER) database is a large database of US patients, but the detailed information regarding disease treatments is not mentioned.
- Competing risk model was performed to eliminate the effect of death, which would lead to a bias for the incidence of sAML.
- Case-control matching analysis was performed to eliminate the effect of confounding variables between sAML group and dnAML group.

Introduction

Diffuse large B-cell lymphoma (DLBCL) is the most common and aggressive type of lymphoma.[1] The combination of rituximab with CHOP (cyclophosphamide, doxorubicin, vincristine and prednisolone) chemotherapy has improved the overall survival (OS) of patients with DLBCL by at least 20%.[2] However, with the increasing survival of patients after DLBCL, the risk of second primary malignancies (SPMs) has also increased, and their management has become an emerging challenge. Currently, SPMs are an important cause of death among survivors of DLBCL.[3,4]

One of the main secondary malignancies following DLBCL is acute myeloid leukemia (AML). For years, more cases of AML have been reported in survivors DLBCL than in the general population.[5,6] Although the underlying factors and biological mechanisms of AML following DLBCL need to be better clarified, the factors about treatment, including the use of rituximab, have been thought to be the main cause of the increased risk.[7-9] The management of patients with secondary AML (sAML) may be challenging because of cumulative toxicity from the treatment of primary DLBCL. Previous exposure to treatment-related factors, including radiotherapy and systemic chemotherapy, has limited the treatment options for secondary neoplasms and further alters their outcomes.[10-13] Hence, we asked if the outcome of secondary AML was poorer than that of de novo AML (dnAML). Meanwhile, considering the difficulty in the management of patients with sAML, a search for the predictive factors for the occurrence of sAML will be meaningful. As far as we know, information available regarding the incidence and prognosis of sAML following DLBCL is limited.

Here, we used sequential cancer data available from the large and high-quality, population-based National Cancer Institute Surveillance, Epidemiology, and End Results (SEER) program to describe the pattern of incidence, investigate the predictive factors for the occurrence of sAML, and compare the outcomes of patients with sAML with their de novo counterparts.

Materials and Methods

Data Source and Sample

The SEER program's research data for 17 registries (excluding Alaska) were used to assess the incidence and explore the hazard factors of sAML in survivors of primary DLBCL diagnosed between 2000 and 2016.[14] DLBCL cases were identified according to the Lymphoma Subtype Recode/WHO 2008, which is updated for hematopoietic conditions and coded based on ICD-O-3 morphology codes (DLBCL: 9678-9680, 9684, 9688, 9712, 9735, and 9737-9738) and the WHO Classification of Tumors of Hematopoietic and Lymphoid Tissues (2008).[15] We excluded cases which were coded as autopsy or death-certificate-only, where DLBCL was not the first primary cancer, and those with unknown age or race. To exclude patients with synchronous DLBCL

and AML, cases diagnosed with AML within the first 2 months of being diagnosed with DLBCL were not included in this study, as well as those with <2 months of follow-up. The process of cases selection was shown in Figure 1. At last, a total of 70280 patients with primary DLBCL were identified, and by the end of the follow-up, 264 of them had developed sAML. For each case, age, gender, year of diagnosis, Ann Arbor stage, survival status, follow-up time, interval time between the diagnosis of DLBCL and sAML, and some other information were extracted from SEER. And, the Ann Arbor stage at diagnosis of DLBCL was classified into two: stage Ⅰ and Ⅱ disease as early stage, and stage Ⅲ and Ⅳ disease as advanced stage.

In order to explore whether there was a difference in survival outcome between patients with sAML and patients with dnAML, we first listed the detailed characteristics of all patients with sAML (two cases with unknown survival time were excluded). We then obtained cases list of dnAML from the SEER database using the same histological subtype as for cases of sAML, but the AML was the first malignancy for a given individual.[14] Finally, a total of 30835 patients were identified from SEER database in 2000-2016.

Statistical Analysis

Kolmogorov-Smirnov normality test was performed to examined the distributions of continuous numerical variables. Variables that did not conform to a normal distribution were described by median and range, and comparison was done with the Wilcoxon rank-sum nonparametric test. Otherwise, data are expressed by means and standard deviations (SDs), and t-test or variance analysis was used for the comparison. Differences in proportions across the groups were compared with the chi-square test. The calculation of the standardized incidence ratio (SIR) and 95% confidence interval (CI) for sAML in patients with DLBCL was performed using SEER*stat software. And, the calculation of person-years for DLBCL survivors was also performed by SEER*stat.

The analyses of cumulative incidence of sAML were completed using competing risk analysis, in which death from any cause was considered the sole competing risk. The differences in cumulative incidence among the groups were compared using Gray's test. Furthermore, to explore the risk factors for sAML, a regression analysis using the semiparametric proportional hazards model proposed by Fine and Gray was performed.[16] Using these models, the semiparametric hazard ratios (sHRs) and their 95% CIs for risk factors were estimated.

To compared the survival outcome of patients with sAML and patients with dnAML, we performed a case-control matching analysis. Based on age (±2 years), calendar year of diagnosis (±2 years), sex, and race, we matched sAML with dnAML patients at a 1:1 ratio. Case matching was completely random and the variables (survival status and cause of death) that might affect the matching result were with no awareness. Because the SEER database does not have detailed information about the treatment, matching

for therapy was impossible. We used a shared-frailty Cox model to interpret the 1:1 matched design. Meanwhile, the factors, age, sex, race, and number of years of diagnosis were adjusted for the model. For AML with previous DLBCL versus dnAML, the hazard ratio (HR) and its 95% CI was calculated.

R software (version 3.6.3) with "cmprsk" and "survival" packages, STATA (version 14.0; Stata Corporation, College Station, TX, USA), and SEER * stat software (version 8.3.6, NCI, NIH, Bethesda, MD, USA) were used to perform these analyses. In this study, we treated a two-sided *P* value < 0.05 to be a statistically significant difference.

Patient and public involvement

No patients or public were involved in this study.

Results

In this study, we identified a total of 70280 primary DLBCL patients, and the median follow-up is 90 months (range, 2–203 months), contributing to a total follow-up of 334,516 person-years. By the end of the follow-up, 264 of these cases having a diagnosis of sAML. The median interval between the diagnosis of DLBCL and sAML was 44 months (range, 3–178 months). The characteristics of the entire cohort of patients with DLBCL who have or have not developed sAML are shown in Table 1.

Table 1. Characteristics of two-month survivors of DLBCL reported to the SEER Program (2000-2016)

Characteristic	All n=70280	No sAML n=70016	With sAML n=264	*P*
Follow-up (range), mo	90 (2-203)	90 (2-203)	157 (3-198)	
Age (range), yrs	64 (0-106)	64 (0-106)	63.5 (12-88)	0.197
Age group, yrs				< 0.001
< 60	28289 (40.3%)	28186 (40.3%)	103 (39.0%)	
60-74	23796 (33.9%)	23677 (33.8%)	119 (45.1%)	
75+	18195 (25.9%)	18153 (25.9%)	42 (15.9%)	
Sex				0.165
Male	38409 (54.7%)	38253 (54.6%)	156 (59.1%)	
Female	31871 (45.3%)	31763 (45.4%)	108 (40.9%)	
Race				0.490
Black	5499 (7.8%)	5483 (7.8%)	16 (6.1%)	
White	58661 (83.5%)	58434 (83.5%)	227 (86.0%)	
Other	6120 (8.7%)	6099 (8.7%)	21 (8.0%)	

Primary site				0.015
Nodal	46241 (65.8%)	46048 (65.8%)	193 (73.1%)	
Extranodal	24039 (34.2%)	23968 (34.2%)	71 (26.9%)	
Ann Arbor stage				0.001
Stage Ⅰ	18535 (26.4%)	18470 (26.4%)	65 (24.6%)	
Stage Ⅱ	13717 (19.5%)	13672 (19.5%)	45 (17.0%)	
Stage Ⅲ	10726 (15.3%)	10672 (15.2%)	54 (20.5%)	
Stage Ⅳ	20026 (28.5%)	19937 (28.5%)	89 (33.7%)	
Unknown	7276 (10.4%)	7265 (10.4%)	11 (4.2%)	
Years of diagnosis				< 0.001
2000-2005	23047 (32.8%)	22933 (32.8%)	114 (43.2%)	
2006-2011	25196 (35.9%)	25084 (35.8%)	112 (42.4%)	
2012-2016	22037 (31.4%)	21999 (31.4%)	38 (14.4%)	

Abbreviations: DLBCL, diffuse large B-cell lymphoma; sAML, secondary acute myeloid leukemia

The SIR for sAML overall was 6.23 (95% CI, 5.50–7.03), indicating an elevated incidence compared with that for the general population of the USA. The forest plot for the SIRs is shown in Figure 2. The SIR was 13.46 (95% CI, 10.99–16.33) in patients aged <60 years, 6.17 (95% CI, 5.11–7.39) in patients aged 60–74 years, and 2.72 (95% CI, 1.96–3.68) in patients aged ≥75 years; thus, it decreased with increasing age (*P* for trend < 0.001). The nodal DLBCL had a higher SIR for sAML than extranodal DLBCL. As for the Ann Arbor Stage of DLBCL, the SIR was less for the early-stage as compared to that for advanced-stage disease. Patients with a latency of 24–59 months had a higher SIR than those with other latencies. For the groups of sex, race, and years of diagnosis, no heterogeneity or trend for SIRs was observed.

Further, when competing causes of deaths were considered, the cumulative incidence of sAML was 0.30% (95% CI, 0.26%–0.35%), 0.53% (95% CI, 0.46%–0.60%), and 0.61% (95% CI, 0.53%–0.70%) at 5, 10, and 15 years after the diagnosis of DLBCL, respectively. Moreover, we found that the cumulative incidence of sAML was closely related to the patients' age at the diagnosis of DLBCL (*P*<0.001), the primary site (*P*=0.010), and the Ann Arbor stage of DLBCL (*P*=0.007). The cumulative incidence at 10 years after DLBCL diagnosis was 0.51% (95% CI, 0.41%–0.62%) in patients aged <60 years, 0.74% (95% CI, 0.61%–0.89%) in patients aged 60–74 years, and 0.29% (95%

CI, 0.21%–0.39%) in patients aged ≥75 years. For extranodal DLBCL, the cumulative incidence in patients at 10 years was 0.40% (95% CI, 0.31%–0.51%); it was 0.59% (95% CI, 0.51%–0.69%) for DLBCL occurring in the lymph node. As regards the Ann Arbor Stage of DLBCL, the cumulative incidence in patients at 10 years was 0.43% (95% CI, 0.35%–0.52%) and 0.66% (95% CI, 0.55%–0.79%) for the early and advanced stages, respectively (Figure 3).

Furthermore, according to the semiparametric proportional hazards model, we investigated the risk factors for sAML occurrence. The results are presented in Table 2. Univariate analyses showed that patients' age, primary site, and Ann Arbor stage of DLBCL were statistically significant risk factors (P<0.05). These three variables were selected for the final multivariate analysis, which showed that the patients' age at diagnosis and the Ann Arbor stage of DLBCL were independent predictors of the occurrence of sAML. Patients aged 60–74 years were more likely to have sAML than those aged <60 years. However, patients aged ≥75 years were less likely to have sAML than patients aged <60 years. Patients with advanced-stage DLBCL were more prone to sAML than those with early-stage disease.

Table 2. Univariate and multivariate analyses for predictive factors of developing sAML.

Factors	Univariate			Multivariate		
	sHR	95% CI	*P*	sHR	95% CI	*P*
Age, yrs						
< 60	ref.			ref.		
60-74	1.421	1.092-1.850	0.009	1.417	1.087-1.850	0.010
75+	0.635	0.443-0.908	0.013	0.648	0.452-0.930	0.018
Sex						
Male	ref.					
Female	0.827	0.647-1.060	0.130			
Race						
White	ref.					
Black	0.764	0.460-1.270	0.300			
Other	0.947	0.605-1.480	0.810			
Primary site						
Nodal	ref.			ref.		
Extranodal	0.703	0.536-0.923	0.011	0.770	0.583-1.020	0.065
Ann Arbor stage						
Early stage	ref.			ref.		
Advanced stage	1.423	1.110-1.830	0.005	1.307	1.012-1.690	0.040

Unknown	0.808	0.434-1.500	0.500	0.807	0.431-1.510	0.500

Abbreviations: sAML, secondary acute myeloid leukemia; sHR, subdistribution hazard ratios; ref, reference

The Table 3 listed the characteristics of patients with sAML and their dnAML counterparts. The median survival time for patients with sAML and dnAML was 7 months (95% CI, 6–9 months) and 13 months (95% CI, 10–17 months), respectively (Figure 4). The Cox model showed that patients with sAML had a higher risk of death and a shorter OS than their dnAML counterparts (HR=1.25; 95% CI, 1.01–1.53; *P*=0.038). Of all the causes of death, AML is the most common in patients with both sAML and dnAML. However, we found that death from DLBCL was still a main component of overall mortality for patients who subsequently developed sAML. (Table 4).

Table 3. The baseline characteristics of patients with sAML and the matched cases with dnAML.

Characteristics	sAML n=262	dnAML n=262
Age (range), yrs	68.0 (15.0-95.0)	67.5 (15.0-93.0)
Age group, yrs		
< 60	67 (25.6%)	71 (27.1%)
60-74	121 (46.2%)	120 (45.8%)
75+	74 (28.2%)	71 (27.1%)
Sex		
Male	155 (59.2%)	155 (59.2%)
Female	107 (40.8%)	107 (40.8%)
Race		
Black	16 (6.1%)	16 (6.1%)
White	225 (85.9%)	225 (85.9%)
Other	21 (8.0%)	21 (8.0%)
Year of diagnosis		
Median (range)	2011 (2001-2016)	2010 (2000-2016)

Abbreviations: sAML, secondary acute myeloid leukemia; dnAML, de novo AML

Table 4. Causes of death in patients with sAML and matched dnAML

Causes of Death	sAML (n)	dnAML (n)
No. of deaths	218	210
AML	105	144
DLBCL	66	0

Other cancer	18	21
Other Hematopoietic and lymphoid tumors	12	17
Solid tumor	6	4
Cardiovascular and cerebrovascular	5	6
Infection	4	6
Other	20	28
NA	0	5

Abbreviations: sAML, secondary acute myeloid leukemia; dnAML, de novo AML; AML, acute myeloid leukemia; NHL, non-Hodgkin lymphoma; NA, not available

Discussion

As far as we know, this is the largest population-based study of sAML in patients with DLBCL. In this study, we observed an increased incidence of sAML among survivors of DLBCL and demonstrated substantial heterogeneity in the occurrence of sAML by age at diagnosis, primary site and Ann Arbor stage of DLBCL. Specifically, we identified that the age at diagnosis and stage of DLBCL were independent risk factors for sAML. We also observed that sAML had a shorter OS than dnAML, and that death from DLBCL was a main component of overall mortality for patients who subsequently developed sAML.

In a population-based study, the SIR of sAML was 4.29 for patients with DLBCL, indicating a higher incidence of sAML in patients with DLBCL than that in the general population,[17] which is consistent with our results. Another large study combined data from 25,089 patients with DLBCL from California and reported 75 cases of sAML.[7] This was 4.39-(pre-rituximab) or 8.70-(post-rituximab) times the number of expected cases from the general population, indicating an increased risk, which was similar to that reported herein.

In this study, we confirmed that the SIR of sAML decreased with an increase in age at diagnosis. However, when competing causes of death were considered, patients aged 60–74 years had the highest cumulative probability of sAML at 10 years of follow-up. This result is consistent with that reported in the papers, which show that sAML tends to occur more commonly in older patients.[18–20] However, this study also showed that patients aged ≥75 years had a lower cumulative incidence than younger patients. Since high-dose chemoradiotherapy has been associated with an elevated risk of sAML,[21] and the elderly are usually not given intensive chemotherapy or radiotherapy due to the comorbidity and functional status, which may lower the risk of sAML.[22]

The link of AML risk with the stage at diagnosis of DLBCL has been not well clarified for DLBCL. A large population-based study indicated that patients with advanced-stage

DLBCL were more likely to develop hematological SPMs, and that the most common histology of hematological SPM was AML.[6] In this study, we also found that patients with advanced-stage DLBCL had a higher SIR than those with early-stage DLBCL (P<0.001). Considering the competing causes of death, we found that the cumulative probability of sAML for patients with advanced-stage DLBCL was higher than that for patients with early-stage at 10 years of follow-up. Furthermore, the advanced-stage DLBCL was identified as an independent risk factor for sAML in our study.

In an experimental research with the analysis of gene expression profiling, a number of genes were differentially expressed in patients with early-stage DLBCL compared to those with advanced-stage DLBCL.[23] Another study suggested that increased late relapses in the early-stage DLBCL compared with advanced-stage DLBCL may be caused by biological differences.[24] However, a report of patients with advanced-stage DLBCL also recognized the risk of late relapse.[25] These reports indicate that early-stage DLBCL has a unique biology compared with advanced-stage DLBCL, which may explain the difference in incidence of sAML partly. On the other hand, according to NCCN (National Comprehensive Cancer Network) guideline, patients with early-stage DLBCL usually receive fewer cycles chemotherapy and is treated with local radiotherapy more often than advanced disease.[26] The difference in treatment may also lead to a lower incidence of sAML for patients with early-stage DLBCL.

Our study found that patients with primary sites in the lymph nodes had a higher SIR and cumulative probability than those with extranodal disease. However, according to the semiparametric proportional hazards model, multivariate analysis showed that the primary sites of DLBCL were not independent risk factors for the progress of sAML. As reported in the literature, patients with early-stage DLBCL are more likely to have extranodal disease.[27] This study also showed similar results (data not shown). Given this finding, it is possible that early-stage DLBCL, which is mainly located in extranodal sites, lowers the SIR or cumulative probability.

In the present study, we compared the survival outcomes of patients with sAML and their dnAML counterparts. The results show that the prognosis is worse for patients diagnosed with sAML after surviving DLBCL than those diagnosed with dnAML in matched cases. Previously studies have indicated that the prior therapy of DLBCL shows a detrimental effect, which has been verified in patients who have developed malignant mesotheliomas, bladder cancer, and kidney cancer.[28–30] The successful treatment of second cancer in patients who survive DLBCL has been affected by many factors, such as limitations on the dose and site of radiotherapy, a poor tolerance to chemotherapy, and impaired physiologic reserve. Another intriguing factor may result from the intrinsically worse biology of sAML, which require more in-depth research.

Since this is a retrospective observational study based on SEER database, there are some limitations for this study. First, we are limited by the extent of the data for some

covariates of interest. And, one of the primary limitations is that we cannot obtain the detailed information regarding disease treatments. Therefore, it is impossible to establish a correlation between DLBCL treatment and the development of sAML. In addition, the therapeutic modalities that could be used to treat patients with sAML are also not mentioned in the database, which limits the exploration of prognosis. Second, the diagnostic standard and classification are not uniform, such as the diagnosis of AML and the classification of DLBCL, which may impact on the conclusions. Third, we have to exclude some cases with unknown characteristics, and this may lead to a bias of the result. However, there are several novel findings shown in this study. These findings may be helpful in future prospective trials for patients with DLBCL.

In conclusion, the current findings suggest that the incidence of AML increases significantly among survivors of DLBCL. Furthermore, we showed that age and Ann Arbor stage of DLBCL at diagnosis are independent risk factors for sAML. We also found that patients with sAML had shorter OS than their dnAML counterparts. These findings will be beneficial for the management of patients with newly diagnosed DLBCL.

Acknowledgment

We thank Editage (www.editage.cn) for their help in English language editing.

Funding

No funding for this study.

Declaration of Interest

The authors declare no potential conflicts of interest.

Patient consent for publication

Not applicable.

Ethics approval

This study used the SEER research database, which is approved by the NIH Ethics Program.

Authors Contributions

Yu Du: Conceptualization and design, collection and assembly of the data, formal

analysis and interpretation, visualization, writing-original draft, and writing-review and editing. **Ying Wang:** Collection and assembly of the data, data curation, visualization, writing-original draft, and writing-review and editing. **Qinlu Li:** Data curation and analysis, writing-review and editing. **Xiaona Chang:** Methodology, supervision, and writing-review. **Heng Zhang:** Supervision and writing-review. **Min Xiao:** Conceptualization, methodology, validation, data curation and writing-review and editing. **Shugang Xing:** Conceptualization, methodology, validation, formal analysis, data curation, supervision, writing-original draft, writing-review and editing.

Data availability statement

Data used in this study are available from the Surveillance, Epidemiology, and End Results (SEER) database (https://seer.cancer.gov).

References

[1]. Swerdlow SH, Campo E, Harris NL, et al, eds. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues*. 4th ed. IARC Press; 2008

[2]. 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;346:235–242. doi: 10.1056/NEJMoa011795

[3]. de Moor JS, Mariotto AB, Parry C, et al. Cancer survivors in the United States: Prevalence across the survivorship trajectory and implications for care. *Cancer Epidemiol Biomarkers Prev*. 2013;22:561-570. doi: 10.1158/1055-9965.EPI-12-1356

[4]. Chinen Y, Tanba K, Takagi R, et al. Second primary malignancy after rituximab-containing immunochemotherapy for diffuse large B cell lymphoma. *Leuk Lymph*. 2020; 61(14):3378-3386. doi: 10.1080/10428194.2020.1811862

[5]. Gangavarapu LS, Surapaneni R. Diffuse large B-cell lymphoma survivorship: Risk of second malignancies in long term survivors. *Blood*. 2015;126(23):3921. doi: 10.1182/blood.V126.23.3921.3921

[6]. Major A, Smith DE, Ghosh D, et al. Risk and subtypes of secondary primary malignancies in diffuse large B-cell lymphoma survivors change over time based on stage at diagnosis. *Cancer*. 2020;126(1):189-201. doi: 10.1002/cncr.32513

[7]. Tao L, Clarke CA, Rosenberg AS, et al. Subsequent primary malignancies after diffuse large B-cell lymphoma in the modern treatment era. *Br J Haematol*. 2017;178(1):72-80. doi: 10.1111/bjh.14638

[8]. Fleury I, Chevret S, Pfreundschuh M, et al. Rituximab and risk of second primary malignancies in patients with non-Hodgkin lymphoma: a systematic review and meta-analysis. *Ann Oncol*. 2016;27(3):390-397. doi: 10.1093/annonc/mdv616

[9]. Brown JR, Yeckes H, Friedberg JW, et al. Increasing incidence of late second malignancies after conditioning with cyclophosphamide and total-body irradiation and autologous bone marrow transplantation for non-Hodgkin's lymphoma. *J Clin Oncol*. 2005;23(10):2208-2214. doi: 10.1200/JCO.2005.05.158

[10]. Milano MT, Li H, Constine LS, et al. Survival after second primary lung cancer: A population-based study of 187 Hodgkin lymphoma patients. *Cancer*. 2011;117(24):5538-5547. doi: 10.1002/cncr.26257

[11]. Milano MT, Li H, Gail MH, et al. Long-term survival among patients with Hodgkin's lymphoma who developed breast cancer: a population-based study. *J Clin Oncol*. 2010;28(34):5088-5096. doi: 10.1200/JCO.2010.29.5683

[12]. Hemminki K, Lenner P, Sundquist J, et al. Risk of subsequent solid tumors after non-Hodgkin's lymphoma: effect of diagnostic age and time since diagnosis. *J Clin Oncol*. 2008;26(11):1850-1857. doi: 10.1200/JCO.2007.14.6068.

[13]. Herr MM, Schonfeld SJ, Dores GM, et al. Risk for malignancies of infectious etiology among adult survivors of specific non-Hodgkin lymphoma subtypes. *Blood Adv*. 2019;3(13):1961-1969. doi: 10.1182/bloodadvances.2019030924

[dataset] [14]. Surveillance, Epidemiology, and End Results (SEER) Program

(www.seer.cancer.gov) SEER*Stat Database: Incidence - SEER 18 Regs excluding AK Research Data, Nov 2018 Sub (2000-2016) <Katrina/Rita Population Adjustment>, National Cancer Institute, DCCPS, Surveillance Research Program, released April 2019, based on the November 2018 submission.

[15]. Sabattini E, Bacci F , Sagramoso C, et al. WHO classification of tumors of haematopoietic and lymphoid tissues in 2008: An overview. *Pathologica*.2008;102(3):83-87.

[16]. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc*. 1999;94(446):496-509.

[17]. Jiang S, Zhen H, Jiang H. Second primary malignancy in diffuse large B-cell lymphoma patients: A SEER database analysis. *Curr Probl Cancer*. 2020;44(1):100502. doi: 10.1016/j.currproblcancer.2019.100502

[18]. Szotkowski T, Rohon P, Zapletalova L, et al. Secondary acute myeloid leukemia - a single center experience. *Neoplasma*. 2010;57(2):170-178. doi: 10.4149/neo_2010_02_170

[19]. Ossenkoppele G, Montesinos P. Challenges in the diagnosis and treatment of secondary acute myeloid leukemia. *Crit Rev Oncol Hematol*. 2019;138:6-13. doi: 10.1016/j.critrevonc.2019.03.003

[20]. Yang D, Fu X, Zhang X, et al. Therapy-related acute myeloid leukemia in patients with lymphoma: A report of four cases and review of the literature. *Oncol Lett*. 2015;10(5):3261-3265. doi: 10.3892/ol.2015.3703

[21]. Metayer C, Curtis RE, Vose J, et al. Myelodysplastic syndrome and acute myeloid leukemia after autotransplantation for lymphoma: a multicenter case-control study. *Blood.* 2003;101(5):2015-2023. doi: 10.1182/blood-2002-04-1261

[22]. Morrison VA, Hamlin P, Soubeyran P, et al. Diffuse large B-cell lymphoma in the elderly: impact of prognosis, comorbidities, geriatric assessment, and supportive care on clinical practice. An International Society of Geriatric Oncology (SIOG) expert position paper. *J Geriatr Oncol*. 2015;6(2):141-52. doi: 10.1016/j.jgo.2014.11.004

[23]. Castillo JJ, Winer ES, Olszewski AJ. 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*. 2014;89(3):310-314. doi: 10.1002/ajh.23638

[24]. Roberts RA, Rimsza LM, Staudt L, et al. Gene expression differences between low and high stage diffuse large B cell lymphoma (DLBCL). *Blood*. 2006;108(11):809. doi: 10.1182/blood.V108.11.809.809

[25]. Stephens DM, Li H, LeBlanc ML, et al. Continued Risk of Relapse Independent of Treatment Modality in Limited-Stage Diffuse Large B-Cell Lymphoma: Final and Long-Term Analysis of Southwest Oncology Group Study S8736. *J Clin Oncol*. 2016;34(25):2997-3004. doi: 10.1200/JCO.2015.65.4582

[26]. Zelenetz AD, Gordon LI, Chang JE, et al. NCCN Guidelines® Insights: B-Cell Lymphomas, Version 5.2021. *J Natl Compr Canc Netw*. 2021;19(11):1218-1230. doi: 10.6004/jnccn.2021.0054

[27]. Coiffier B, Thieblemont C, Van Den Neste E, et al. Long-term outcome of patients

in the LNH-98.5 trial, the first randomized study comparing rituximab-CHOP to standard CHOP chemotherapy in DLBCL patients: a study by the Groupe d'Etudes des Lymphomes de l'Adulte. *Blood*. 2010;116(12):2040-2045. doi: 10.1182/blood-2010-03-276246

[28]. Chirieac LR, Barletta JA, Yeap BY, et al. Clinicopathologic characteristics of malignant mesotheliomas arising in patients with a history of radiation for Hodgkin and non-Hodgkin lymphoma. *J Clin Oncol*. 2013;31(36):4544-4549. doi: 10.1200/JCO.2013.49.9616

[29]. Travis LB, Curtis RE, Glimelius B, et al. Bladder and kidney cancer following cyclophosphamide therapy for non-Hodgkin's lymphoma. *J Natl Cancer Inst*. 1995;87(7):524-530. doi: 10.1093/jnci/87.7.524

[30]. André M, Mounier N, Leleu X, et al. Second cancers and late toxicities after treatment of aggressive non-Hodgkin lymphoma with the ACVBP regimen: a GELA cohort study on 2837 patients. *Blood*. 2004;103(4):1222-1228. doi: 10.1182/blood-2003-04-1124

Figure Legends:

Figure 1. The process of cases selection.

Figure 2. The standardized incidence ratio (SIR) forest plot for patients with secondary acute myeloid leukemia among survivors of diffuse large B-cell lymphoma. 95% CI, 95% confidence interval; ref, reference.

Figure 3. Cumulative incidence of secondary acute myeloid leukemia among survivors of diffuse large B-cell lymphoma. AML, acute myeloid leukemia.

Figure 4. The comparative outcome between survivors of diffuse large B-cell lymphoma who developed secondary acute myeloid leukemia (sAML) and matching patients with de novo acute myeloid leukemia (dnAML).

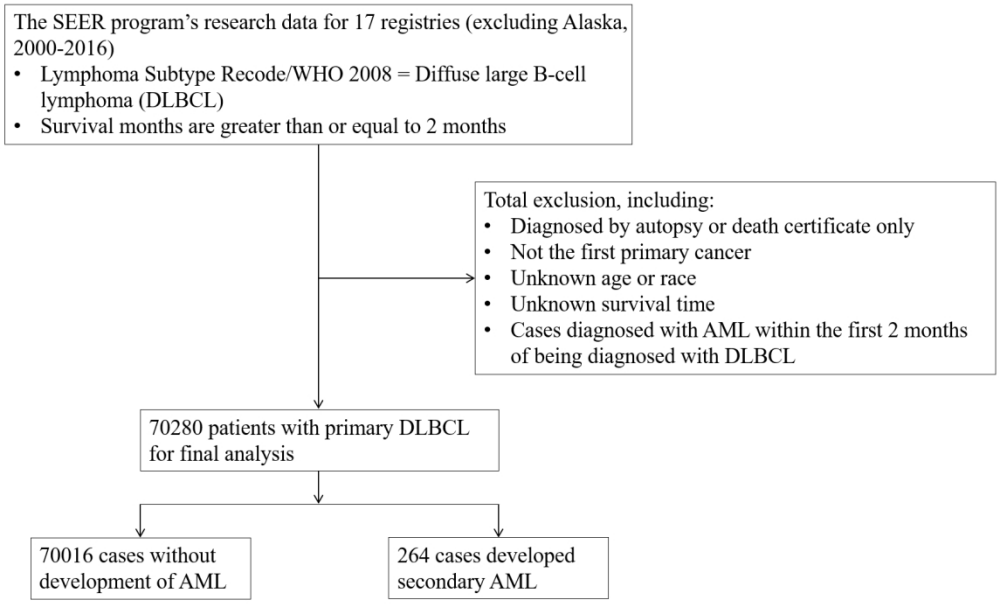

Figure 1. The process of cases selection.

772x466mm (57 x 57 DPI)

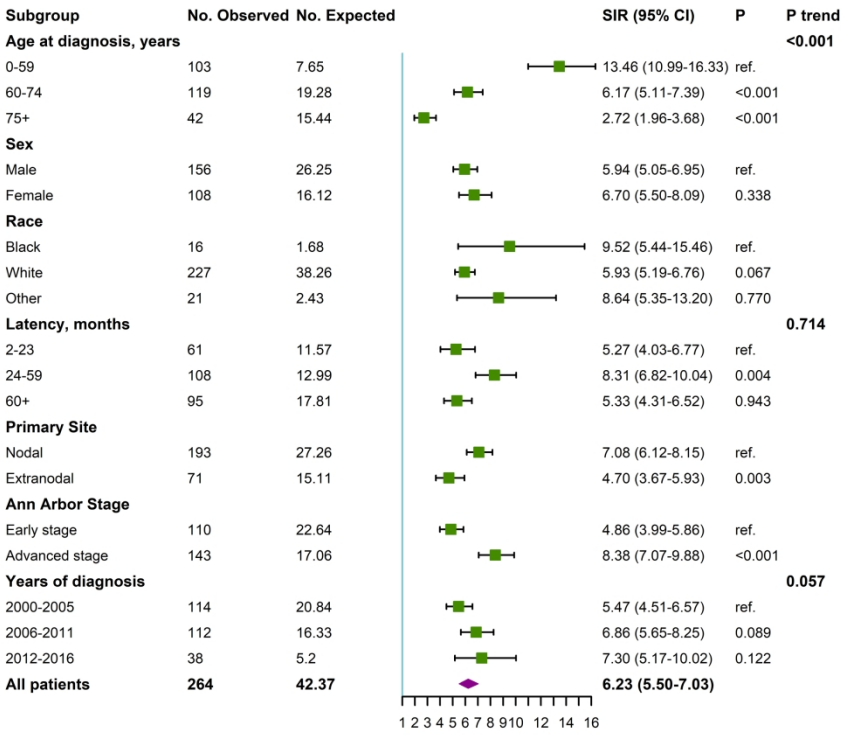

Figure 2. The standardized incidence ratio (SIR) forest plot for patients with secondary acute myeloid leukemia among survivors of diffuse large B-cell lymphoma. 95% CI, 95% confidence interval; ref, reference.

254x217mm (300 x 300 DPI)

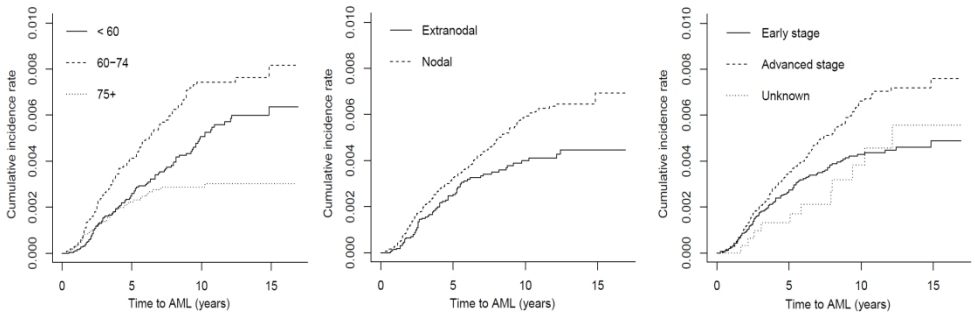

Figure 3. Cumulative incidence of secondary acute myeloid leukemia among survivors of diffuse large B-cell lymphoma. AML, acute myeloid leukemia.

621x241mm (150 x 150 DPI)

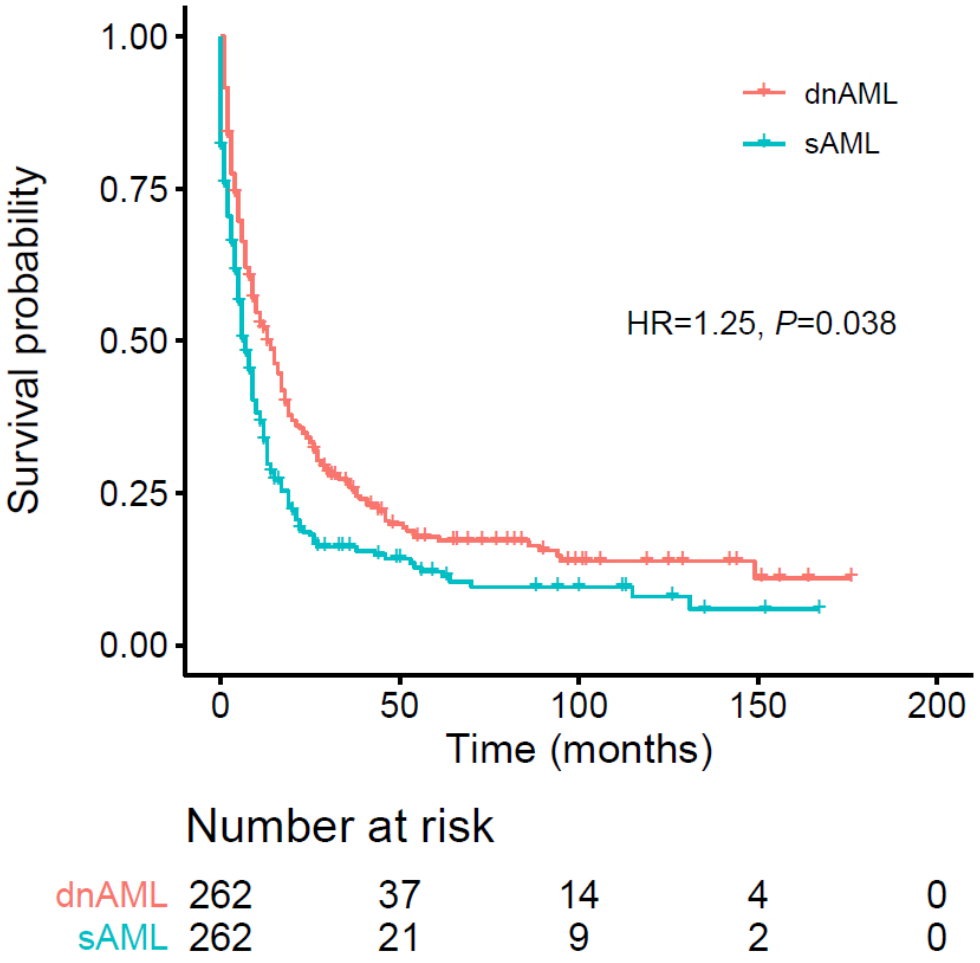

Figure 4. The comparative outcome between survivors of diffuse large B-cell lymphoma who developed secondary acute myeloid leukemia (sAML) and matching patients with de novo acute myeloid leukemia (dnAML).

STROBE Statement—Checklist of items that should be included in reports of ***case-control studies***

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1-2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	1-2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3
Objectives	3	State specific objectives, including any prespecified hypotheses	3
Methods			
Study design	4	Present key elements of study design early in the paper	3
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	3-4
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls	4
		(*b*) For matched studies, give matching criteria and the number of controls per case	4
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	4
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	4
Bias	9	Describe any efforts to address potential sources of bias	4
Study size	10	Explain how the study size was arrived at	4
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	4
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	4
		(*b*) Describe any methods used to examine subgroups and interactions	4-5
		(*c*) Explain how missing data were addressed	4
		(*d*) If applicable, explain how matching of cases and controls was addressed	4
		(*e*) Describe any sensitivity analyses	
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	5
		(b) Give reasons for non-participation at each stage	5
		(c) Consider use of a flow diagram	5
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	5
		(b) Indicate number of participants with missing data for each variable of interest	5
Outcome data	15*	Report numbers in each exposure category, or summary measures of exposure	6

Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	6
		(*b*) Report category boundaries when continuous variables were categorized	6
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	6
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	7
Discussion			
Key results	18	Summarise key results with reference to study objectives	8-9
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	10
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	9
Generalisability	21	Discuss the generalisability (external validity) of the study results	9
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	11

*Give information separately for cases and controls.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.