International Journal of Community Medicine and Public Health
Nnawuba K et al. Int J Community Med Public Health. 2024 Jul;11(7):2533-2539
http://www.ijcmph.com
pISSN 2394-6032 | eISSN 2394-6040

DOI: https://dx.doi.org/10.18203/2394-6040.ijcmph20241805

Original Research Article

Analyzing the incidence, survival, and demographic trends of B-cell prolymphocytic leukemia

Kingsley Nnawuba[1]*, Ifeanyi Uche[2], Samantha Robinson[3], Obed Asare[4], Tochukwu Nzeako[5], Anup Kumar Trikannad Ashwini Kumar[6], Hanna Jensen[7]

[1]Department of Internal Medicine, University of Arkansas for Medical Sciences, Northwest Regional Campus, Fayetteville, AR, United States
[2]Department of Internal Medicine, Brookdale University Hospital Medical Center, Brooklyn, New York
[3]Department of Mathematics, Fulbright College of Arts and Sciences, University of Arkansas, Fayetteville, AR, United States
[4]Department of Information Systems, Walton College of Business, University of Arkansas, Fayetteville, AR, United States
[5]Department of Internal Medicine, Christiana Care Hospital, Newark, Delaware
[6]Division of Hematology Oncology and Myeloma, University of Arkansas for Medical Sciences, AR, United States
[7]Department of Surgery, University of Arkansas for Medical Sciences, Northwest Regional Campus, Fayetteville, AR, United States

Received: 02 June 2024
Revised: 15 June 2024
Accepted: 17 June 2024

***Correspondence:**
Dr. Kingsley Nnawuba,
E-mail: Kcnnawuba@uams.edu

ABSTRACT

Background: B-cell prolymphocytic leukemia (B-PLL) is a rare disease, consisting <1% of mature B-cell malignancies. B-PLL is often refractory to chemotherapy, with a median survival of 3 years. Due to its rarity, no large cohort studies exist elucidating outcomes.
Methods: All B-PLL patients >15 years were identified in the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) database between 2000-2020. Statistical analysis explored demographic variables like age; categorized as adolescent or young adult (AYA) and adults. In adults, differences in survival due to factors such as sex, race/ethnicity, household income, rurality, and age categorized in 10-15 year buckets was analyzed.
Results: B-PLL patients were predominantly white (78%), over 40 years (96%) and mostly residing in metropolitan areas (90%). Interestingly, the AYA cohort were mostly female (70%). 35% of the AYA patients were Hispanic, while being only 9% in adults. Among adults, the rate of Asian/Pacific-Islander patients that were alive at the time of the data query was 53% compared to 34% in Hispanic, 31% in black, and 24% in white patients ($p=0.025$). Younger age was also associated with greater chances of survival ($p<0.001$).
Conclusions: In line with known poor prognosis of the disease, 23% patients were alive at the time of data query. Female and Hispanic patients were overrepresented in the AYA age group. In the adult group Asian/Pacific-Islander patients had better survival outcomes, as did younger patients. Further research is necessary to explore why B-PLL incidence in AYA patients is higher among Hispanic and females.

Keywords: Hematological malignancies, Leukemia, Prolymphocytic leukemia, Cancer survival

INTRODUCTION

B-cell prolymphocytic leukemia (B-PLL) is a rare hematologic malignancy, constituting less than 1% of mature B-cell neoplasms.[1] Typically affecting elderly individuals, B-PLL is characterized by peripheral blood prolymphocyte levels exceeding 55%, minimal lymphadenopathy, marked splenomegaly, and markedly elevated white blood cell counts.[2] The disease often exhibits resistance to standard chemotherapy, leading to a generally poor prognosis, with a median survival of approximately three years.[3]

B-PLL predominantly manifests in older adults, typically presenting between ages 65 and 70 years. Both genders are equally affected, although there appears to be a higher prevalence among Caucasians.[4] Initially identified in 1974 by Galton et al, B-PLL was formally recognized as a distinct entity by the World Health Organization in 2008, owing to its clinical and pathological similarities to chronic lymphocytic leukemia (CLL).[3]

Due to its rarity, there are limited cohort studies examining survival patterns among B-PLL patients. This study aims to delineate the demographic characteristics of B-PLL patients using the Surveillance, Epidemiology, and End Results (SEER) database from the National Cancer Institute, with a focus on factors potentially influencing survival trends.

METHODS

This retrospective descriptive study was deemed exempt by the Institutional Review Board. All patients confirmed to have B-PLL between the years 2000 and 2020 were retrieved from The Surveillance, Epidemiology, and End Results (SEER) national database, which provides deidentified but comprehensive national-level clinical data for investigation.[5]

The inclusion criteria and clinical variables extracted for this study were: B-cell prolymphocytic leukemia (ICD code 16.32, site recode, international classification of diseases for oncology (ICD-O-3)/WHO 2008), and age ≥15. The demographic characteristics queried included sex race and ethnicity, median household income, and the rurality of the patient's residence. Survival status was considered as the main outcome variable. The patients were categorized as either adult (>40 years of age) or adolescent or young adult (AYA), if between ages 15 and 39 years. However, the AYA patients were excluded from inferential analysis due to their extremely small sample size. Exclusion criteria for participation in this study included; age <15, patients with wrong or no diagnosis of the BPLL, incomplete treatment data, missing demographic information, non-primary cases and lack of living vs dead status.

Descriptive statistics were calculated for both the full and analytic samples and were also calculated to reveal possible demographic differences between adult and AYA cohorts. Bivariate associations between living status and each demographic characteristic were explored utilizing chi-square tests. To evaluate the unique contributions of each demographic characteristic on survival status while controlling for other factors, multivariable logistic regression and conditional inference tree analyses were conducted. Conditional inference trees are nonparametric tree-based models that utilize recursive partitioning to identify the most influential factors contributing to living status as determined by permutation-based significance testing.[6] This additional multivariable analysis is appropriate even when model assumptions for traditional parametric statistical are violated and is distinctive in its ability to detect and explore complex interaction effects that need not be specified a priori. All data pre-processing and analyses were performed in R Version 4.3.3 (R Core Team (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/) with statistical significance defined as $p<0.05$.

RESULTS

Overall patient characteristics can be found summarized in Table 1. This was a predominantly male (62%) and mostly white (78%) cohort, with most patients (96%) over 40 years of age, living in very large metropolitan areas (90%), including a considerable portion of them living in areas with populations exceeding one million individuals (61%). In this 20-year cohort, 23% of patients were alive at the time of the data query.

Table 1: Sample characteristics.

Demographic	Full sample (%) (n=608)	Analytic sample (%) (n=449)
Sex		
Male	375 (61.7)	278 (61.9)
Female	233 (38.3)	171 (38.1)
Race/ethnicity		
Hispanic (all races)	60 (9.9)	41 (9.1)
Non-Hispanic Asian or Pacific Islander	26 (4.3)	19 (4.2)
Non-Hispanic Black	47 (7.7)	35 (7.8)
Non-Hispanic White	475 (78.1)	354 (78.8)

Continued.

Demographic	Full sample (%) (n=608)	Analytic sample (%) (n=449)
Median household Income (adjusted for inflation)		
<$45 K	20 (3.3)	17 (3.8)
≥$45 K	588 (96.7)	432 (96.2)
Rural-urban continuum		
Metropolitan areas w/pop. ≥1 million	373 (61.3)	267 (59.5)
Metropolitan areas w/pop. Between 250 K and 1 million	124 (20.4)	91 (20.3)
Metropolitan areas w/ pop. <250 K	51 (8.4)	42 (9.4)
Nonmetropolitan areas adjacent to a metropolitan area	44 (7.2)	36 (8.0)
Nonmetropolitan areas not adjacent to a metropolitan area	16 (2.6)	13 (2.9)
Age group (years)		
Adolescent and young adult (AYA) i.e., ≤39	23 (3.8)	NA*
40-54	49 (8.1)	45 (10.0)
55-64	98 (16.1)	80 (17.8)
65-74	159 (26.1)	119 (26.5)
75-84	193 (31.7)	148 (33.0)
85 or above	86 (14.1)	57 (12.7)
Living status		
Alive	139 (22.9)	120 (26.7)
Dead	332 (54.6)	329 (73.3)
Dead (due to causes other than this cancer)	137 (22.5)	NA

*NA - Not Applicable.

Table 2: Demographics by age group.

Demographic	AYA (%) (n=23)	Adult (%) (n=585)
Sex		
Male	7 (30.4)	368 (62.9)
Female	16 (69.6)	217 (37.1)
Race/ethnicity		
Hispanic (all races)	8 (34.8)	52 (8.9)
Non-Hispanic Asian or Pacific Islander	2 (8.7)	24 (4.1)
Non-Hispanic Black	2 (8.7)	45 (7.7)
Non-Hispanic White	11 (47.8)	464 (79.3)
Median household income (adjusted for inflation)		
<$45 K	1 (4.3)	19 (3.2)
≥$45 K	22 (95.7)	566 (96.8)
Rural-urban continuum		
Metropolitan areas w/pop. ≥1 million	15 (65.2)	358 (61.2)
Metropolitan areas w/pop. between 250 k and 1 million	3 (13.0)	121 (20.7)
Metropolitan areas w/pop. 250 k	4 (17.4)	47 (8.0)
Nonmetropolitan areas adjacent to a metropolitan area	0 (0.0)	44 (7.5)
Nonmetropolitan areas not adjacent to a metropolitan area	1 (4.3)	15 (2.6)
Living status		
Alive	19 (82.6)	120 (20.5)
Dead	3 (13.0)	329 (56.2)
Dead (due to causes other than this cancer)	1 (4.3)	136 (23.2)

Table 2 highlights the descriptive differences in demographics between adult and AYA patient groups. Female patients were overrepresented in the AYA cohort, whereas the adult cohort was predominantly male. While both age groups were predominantly white, the AYA group had significantly more hispanic patients (35%) compared to the adult cohort (9%). As expected, the AYA had more alive patients at the time of the data query (83%) compared to the adult cohort (21%). Significance of such differences was not assessed due to the small AYA sample size.

Table 3: Bivariate analysis of demographics by living status.

Demographic	Alive (%)	Dead (%)	P value[1]
Sex			1.000
Male (n=278)	74 (26.6)[2]	204 (73.4)	
Female (n=171)	46 (26.9)	125 (73.1)	
Race/ethnicity			0.025
Hispanic (all races) (n=41)	14 (34.1)	27 (65.9)	
Non-Hispanic Asian or Pacific Islander (n=19)	10 (52.6)	9 (47.4)	
Non-Hispanic Black (n=35)	11 (31.4)	24 (68.6)	
Non-Hispanic White (n=354)	85 (24.0)	269 (76.0)	
Median household income (adjusted for inflation)			0.797
<$45 K (n=17)	4 (23.5)	13 (76.5)	
≥$45 K (n=432)	116 (26.9)	316 (73.1)	
Rural-urban continuum			0.093
Metropolitan areas w/ pop. ≥1 million (n=267)	78 (29.2)	189 (70.8)	
Metropolitan areas w/ pop. between 250 k and 1 million (n=91)	28 (30.8)	63 (69.2)	
Metropolitan areas w/ pop. <250 k (n=42)	6 (14.3)	36 (85.7)	
Nonmetropolitan areas adjacent to a metropolitan area (n=36)	5 (13.9)	31 (86.1)	
Nonmetropolitan areas not adjacent to a metropolitan area (n=13)	3 (23.1)	10 (76.9)	
Age group (years)			<0.001
40-54 (n=45)	22 (48.9)	23 (51.1)	
55-64 (n=80)	25 (31.3)	55 (68.8)	
65-74 (n=119)	43 (36.1)	76 (63.9)	
75-84 (n=148)	23 (15.5)	125 (84.5)	
85 or above (n=57)	7 (12.3)	50 (87.7)	

[1]P value for demographic survival differences utilizing a $\chi 2$ test if assumptions are met otherwise the p value for demographic differences is computed for the $\chi 2$ test using a Monte Carlo test with 10,000 replicates. [2]Values throughout reflect N (%) of each row variable that are alive and dead respectively e.g., 74 individuals among the 278 males (26.6%) are alive.

Bivariate analyses

AYA patients and those patients classified as dead due to causes other than this cancer were removed from the analytic sample for bivariate and multivariable analyses. Table 3 focuses this analytic sample of only adult patients, outlining the differences in living status that exist within this cohort based due to demographics. Living status differed significantly based upon age group (p<0.001) and race/ethnicity (p=0.025).

Multivariable analyses

Multivariable analyses using both a tree-based model and multivariable logistic regression were conducted to determine how different features impacted survival, while accounting for other variables.

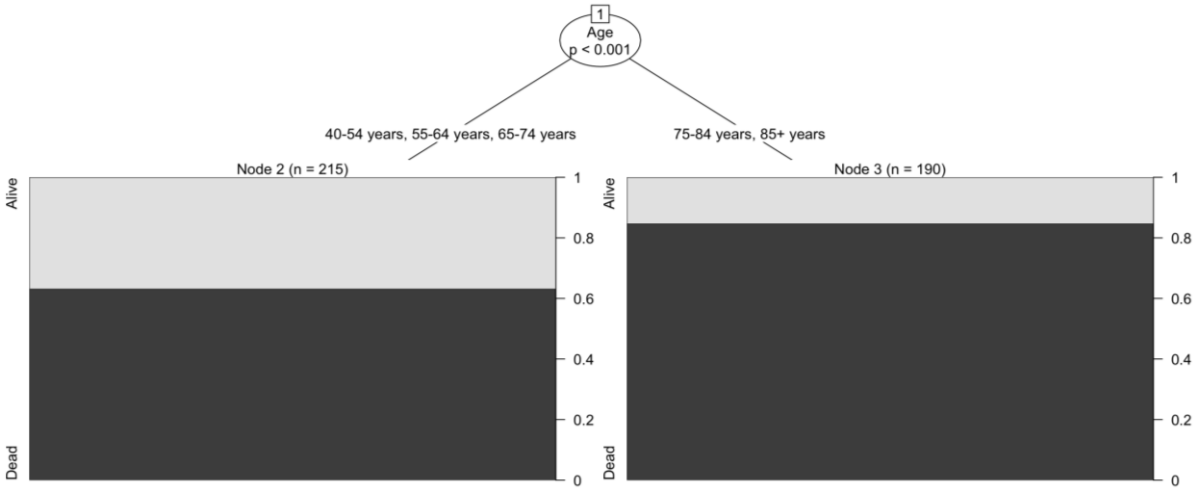

Figure 1: Tree-based model.

A conditional inference tree was constructed from all the variables to determine what was most predictive of current living status (Alive or Dead). The below is the result of the overall model.

This model was based on using 90% of the sample as a training set and 10% of the sample as a testing set, resulting in approximately 73% accuracy (95% CI: [57.2, 85.0]) in predicting whether a patient was currently alive. The variable that emerged as most significant was age, with those in age groups below 75 years of age much more likely to be currently alive. All variables not included in the above figure were not statistically significant.

Multivariable logistic regression

We fitted a multivariable logistic model to predict whether a patient was currently alive using age, sex, race, median household income, and rurality as explanatory variables. Standardized parameters were obtained by fitting the model on a standardized version of the dataset and p values were computed using a Wald z-distribution approximation.

Results of the model indicated that the effect of age was the only variable that was statistically significant ($p<0.001$). Table 4 below provides odds ratios from the logistic regression model.

Table 4: Multivariable logistic regression results.

Demographic	Odds ratio	P value
Intercept	0.94	0.949
Sex		
Male	1.24	0.370
Race/ethnicity		
Non-Hispanic Asian or Pacific Islander	0.50	0.246
Non-Hispanic Black	1.52	0.427
Non-Hispanic White	1.32	0.452
Median household income		
≥$45 K	1.84	0.416
Rural-urban continuum		
Metropolitan w/pop. ≥1 million	0.38	0.046*
Metropolitan w/pop. 250 K to 1 million	0.31	0.023*
Nonmetropolitan, metropolitan adjacent	0.90	0.885
Nonmetropolitan, not metropolitan adjacent	0.73	0.741
Age group (years)		
55-64	2.13	0.062
65-74	1.72	0.145
75-84	5.29	<0.001*
85 or above	7.84	<0.001*

*statistically signinficant.

In the model, increasing age was significantly associated with poorer prognosis. When looking at the odds ratios and comparing individuals with the same sex, race, median household income, and rurality, the odds of a 75 to 84 year-old individual being dead are over 5 times that of a 45-54 year-old individual (OR=5.29, $p<0.001$) and the odds of an individual 85 years-old or above not being alive are almost 8 times that of a 45-54 year-old individual (OR=7.84, $p<0.001$). Larger metropolitan areas of residence seemed to be associated with improved survival.

DISCUSSION

In this study, mortality outcomes and predictors of mortality were examined in patients with B-PLL. Non-Hispanic Whites had the highest mortality rate (76.0%). More females were diagnosed among adolescents and young adults (≤39 years) than males. The incidence and mortality of B-PLL were higher in persons over the age of 40 than in AYA. In addition, young adults and adolescents were more likely to be Hispanic than was seen in the adult cohort. Overall, age was the strongest predictor of survival in the adult cohort with younger aged patients having a higher chance of being listed as currently alive.

At the end of the study period, 55% of the overall sample (n=332) had died from B-PLL with only 22.9% (n=139) still alive at the time of the data query. This is in keeping with other studies that outline the poor prognosis in B-PLL patients.[7] We identified various factors that could be associated with this increased risk of mortality. Non-Hispanic Whites tend to have more deaths amongst their racial group from B-PLL compared to other age groups, with only 76% of Non-Hispanic Whites in the analytical sample having died from B-PLL. Other studies have identified Caucasian race as a predictor of poor survival in B-PLL.[8]

From our observation non-Hispanic Asian or Pacific Islanders had a better outcome compared to other races

with 53% of the analytic sample alive at the end of the study period. Hispanics and Non-Hispanic Blacks followed with 34% and 31% alive at the conclusion of the study respectively. Non-Hispanic Whites (NHW) had poorer outcomes with only 24% alive at the end of the study period.

The gender distribution of B-PLL among the population studied was interesting. We found that amongst adolescents and young adults (AYA), those aged less than 39 years, there were more females compared to males who were diagnosed with B-PLL (70% vs 30%). However, amongst adults, those aged >40 years we found the opposite with men more than women (63% vs 37%). From other studies there appears to be a small increase in the incidence in males compared to females (n=2) while another study identified an equal distribution.[4]

Anecdotally, when we compared AYA to older adults, we observed differences in the distribution among race groups. Among the AYA there were more Hispanic patients with the disease when compared to older adults. We also observed a smaller proportion of NHWs in the AYA group with only 47.8% being NHW compared to 79.3% in the older adult group. Studies have shown that NHWs had a significantly higher incidence of B-PLL when compared to other races.[7] From our study AYA who identify as NHW made up less than half of the population, an interesting and unexpected finding. However, due to the extremely small sample size of AYA patients in this cohort, we are unable to conclude whether these trends were statistically significant.

It has been shown that among people under the age of 20, Hispanics have the highest rate of leukemias of all ethnicities—in children, by more than 20% according to cancer organizations.[9] Researchers have found genetic and genomic mutations linked to ALL in people of Latino descent, although its prevalence in the Latino community could also be environmental.[10] Risk factors may include certain jobs such as farmers involved in cattle ranching, dairy production and herbicide use and exposure to some pesticides.

In keeping with other studies we observed that B-PLL was more predominant in individuals aged >40 years when compared to AYAs.[3,11] Older patients were also more likely to die from this disease and this was not a surprise as the median age of presentation is 69 years with an average survival of about 65 months.[11] From our analysis we concluded that the most significant predictor of survival was age with those adults in younger age cohorts to be more likely to survive compared to older adults (Figure 1). From other studies, other predictors of survival include anemia, TP53 mutation and lymphocyte count, however these variables were not included in our study as they are not tracked in the SEER database.[3]

In this study, aside age and race, residential location seemed to have an association with survival outcomes of patients. Data analysis highlighted that more patients in the studied populations resided in metropolitan areas, than those in rural settings. Considering the multivariable regression results from table in our results, it appears that patients residing in metropolitan areas seemed to have a better outcome given their lower odds ratio which was also statistically significant. Thus, fair and impactful targeted interventions to increase access to care and improve outcomes, particularly in underserved areas, are necessary.

Limitations

Several limitations exist in this SEER-based analysis. One of such is the absence of in-depth clinical and genetic information necessary to understand B-PLL's pathogenesis and therapeutic effectiveness within the sampled population. Equally, there are also few treatment details and compliance data that may perhaps influence on survival analysis. There can be coding as well as reporting errors resulting into misclassification; these usually get compounded when not all B-PLL cases have been captured within the database hence underestimating their incidence rates. Inadequate follow-up periods and possibility for selection bias further complicate them. The most important, however, is that demographic data from SEER may miss out crucial factors such as socioeconomic status and comorbidities. Finally, this is a retrospective study with all the inherent limitations of such research.

CONCLUSION

This retrospective study explored the demographics and outcomes of patients diagnosed with B-cell prolymphocytic leukemia (B-PLL) in the Surveillance, Epidemiology, and End Results (SEER) national database. Overall, we observed a predominantly male (62%) and mostly white (78%) cohort, with most patients (96%) over 40 years of age, living in large metropolitan areas (90%), including a considerable portion of them living in areas with populations exceeding one million individuals (61%). Female and Hispanic patients were overrepresented in patients under 40 years of age. In this 20 year cohort, 23% of patients were alive at the time of the data query. Non-Hispanic White patients had the poorest survival outcomes overall. In a multivariate model, increasing age was associated with poorer survival outcomes, whereas living in a large metropolitan area was associated with improved survival. Further studies are necessitated to improve care and outcomes for this rare disease.

Funding: No funding sources
Conflict of interest: None declared
Ethical approval: The study was approved by the Institutional Ethics Committee

REFERENCES

1. van der Velden VH, Hoogeveen PG, de Ridder D, Schindler-van der Struijk M, van Zelm MC, Sanders M, et al. B-cell prolymphocytic leukemia: a specific

subgroup of mantle cell lymphoma. Blood. 2014;124(3):412-9.
2. Hercher C, Robain M, Davi F, Garand R, Flandrin G, Valensi F, et al. A multicentric study of 41 cases of B-prolymphocytic leukemia: two evolutive forms. Leuk Lymphoma. 2001;42(5):981-7.
3. Dearden C. How I treat prolymphocytic leukemia. Blood. 2012;120(3):538-51.
4. Yamamoto JF, Goodman MT. Patterns of leukemia incidence in the United States by subtype and demographic characteristics, 1997-2002. Cancer Causes Control. 2008;19(4):379-90.
5. Hayat MJ, Howlader N, Reichman ME, Edwards BK. Cancer statistics, trends, and multiple primary cancer analyses from the Surveillance, Epidemiology, and End Results (SEER) Program. Oncologist. 2007;12(1):20-37.
6. Hothorn T, Hornik K, Zeileis A. Unbiased Recursive Partitioning: A Conditional Inference Framework. Journal of Computational and Graphical Statistics. 2006;15(3):651-74.
7. Ermann DA, Fitzgerald LA, Shah H, Hu B, Stephens DM. Treatment and Overall Survival Outcomes of Patients with B-Cell Prolymphocytic Leukemia. American Society of Hematology Annual Meeting; 2023; San Diego, CA: Blood; 2022: 6668-6669.
8. Ionescu FAN, Khoury J, Gaikazian S, Stender M, Jaiyesimi I. Demographic and Survival in B-Cell Prolymphocytic Leukemia: An Analysis of SEER data. Lymphoma and Myeloma Congress; October 23-26; New York, NY: Journal of Hematology; 2019. p. S27-S8.
9. Miller KD, Ortiz AP, Pinheiro PS, Bandi P, Minihan A, Fuchs HE, et al. Cancer statistics for the US Hispanic/Latino population, 2021. CA Cancer J Clin. 2021;71(6):466-87.
10. Bencomo-Alvarez AE, Rubio AJ, Gonzalez MA, Eiring AM. Blood cancer health disparities in the United States Hispanic population. Cold Spring Harb Mol Case Stud. 2021;7(2).
11. Shvidel L, Shtalrid M, Bassous L, Klepfish A, Vorst E, Berrebi A. B-cell prolymphocytic leukemia: a survey of 35 patients emphasizing heterogeneity, prognostic factors and evidence for a group with an indolent course. Leuk Lymphoma. 1999;33(1-2):169-79.

Cite this article as: Nnawuba K, Uche I, Robinson S, Asare O, Nzeako T, Kumar AKTA, et al. Analyzing the incidence, survival, and demographic trends of B-cell prolymphocytic leukemia. Int J Community Med Public Health 2024;11:2533-9.