

Determining factors contributing to the five-year survival of children suffering from acute lymphoblastic leukemia based on tree survival model in the presence of competing risks

Habibollah Esmaily¹, Somaye Barzanouni^{*2}, Hamid Farhangi³

1. Social Determinants of Health Research Center, Mashhad University of Medical Sciences, Mashhad, Iran.
2. Faculty of Health, Mashhad University of Medical Sciences, Mashhad, Iran.
3. Department of Pediatric Hematology-Oncology, Dr. Sheikh Pediatric Hospital, Mashhad University of Medical Sciences, Mashhad, Iran.

Article Type:
Original Article

Article History:
Received: 01 Jul. 2018
Revised: 27 Sep. 2018
Accepted: 13 Oct. 2018

***Correspondence:**

Somaye Barzanouni

Graduate student of
Biostatistics, Faculty
of Health, Mashhad
University of
Medical Sciences,
Mashhad, Iran.

Email:
s.barzanouni@gmail.com



Abstract

Background and objectives: Leukemia is one of the most prevalent cancers worldwide. The relapse of the disease mitigates patient survival time. The convenience of explaining the results obtained from analyzing tree models have encouraged doctors and paramedics to employ them in their research. The current study is an attempt to determine the five-year survival time and factors influencing it in children suffering from acute lymphoblastic leukemia based on tree survival model in the presence of competing risks.

Methods: The required data were collected from 255 children younger than 15, who suffered from acute lymphoblastic leukemia and referred to Dr. Sheikh Hospital in Mashhad, Iran during the years 2007-2015. Afterwards, the survival of the patient until the end of March 2015 was scrutinized. In this regard, various variables like sex, age, treatment period, white blood cells count, hemoglobin, platelet count, LDH level, CNS involvement, mediastan mass, rheumatologic symptoms, etc. were also considered. The relapse of the disease was considered the desired event, whereas the relapse-free death is called competing event. The survival time of the patients from diagnosis date to the date of event (censoring) was calculated on a monthly basis. The fitting of the model is implemented according to maximum within-node homogeneity, which, in turn, is based on the partition function of sum of squares of Event-Specific Martingale Residual changes.

Results: The estimated mean survival time during the relapse and relapse-free death periods as well as in the presence of either events was obtained 55.51, 47.53 and 44.20 months, respectively, implying a decrease in the mean survival time in the presence of competing risks. White blood cell count and platelet count were considered the most influential factors contributing to the relapse or survival. Three sub-groups of patients at risk were identified, and those with white blood cells ≥ 50000 were recognized as the ones with the least mean survival time.

Conclusion: The factors affecting the survival rate of patients and their spots in the model can be employed in making clinical decisions and proposing therapeutic protocols. Identification of sub-groups with identical mean survival rate is the most salient capabilities of the research model.

Keywords: Decision tree; competing risks; acute lymphoblastic leukemia

Introduction

Cancer is recognized as the second cause of mortality in some parts of the world (1, 2). In Iran, it is the third leading cause of death following cardiovascular diseases and accident. Statistically and epidemiologically speaking, childhood cancer is a general term used to describe a range of cancer types found in children under the age of 15. Accordingly, leukemia is the most predominant cancer ever known among children. Although cancer is rarely seen among children aged 0-14 years old (3-9) so that it accounts for less than 1 percent of the whole cancer types (5-7), leukemia is documented to constitute approximately 32 percent of childhood cancers (10). In the U.S, following accident, leukemia is reported to be the second factor leading to death among children between 0 and 14 years old (11-14). Regarding their cellular origin, leukemia can be divided into myeloid and lymphoid. They have been traditionally designated as acute and chronic based on their clinical course. Acute lymphoblastic leukemia is the most common type of childhood leukemia (15) so that it represents 80 percent of childhood leukemia in developed countries (16).

Along similar lines, MAHAK (society to support children suffering from cancer) conducted a study on the prevalence of cancer in Iran and argued that the incidence rate of cancer has increased by 9 per 100 children during various years so that this figure amounted to 15 per 100 children in 2008 (17). Despite appreciable breakthrough achieved in treating ALL among children within the past 30 years, it is known as the most prevalent type of childhood cancers, and children who have experienced the cancer relapse accounts for the vast majority of children suffering from cancer (18). Treating patients with

recurrent ALL is considerably more arduous in comparison with newly diagnosed cases such that their survival rate is significantly lower than that of the second group (19-20). Medical expenses, hospitalization period and mental disorders all point to the socio-economic burden associated with this cancer (21). Since this cancer poses a real threat to the health status of the younger population of the country, serious attention must be paid to childhood cancers.

Medical research mostly aims at investigating time to event data for death or relapse of a disease which is generally called survival time. Survival analysis involves a set of statistical methods whereby survival data are analyzed (22). While analyzing survival data, one may encounter one or more events, whereas he experiences only one of them during the follow-up period. These various events are called competing risks. Therefore, the analysis of the data suggested two causes of failure that compete to occur. Devising an appropriate model for survival data in the presence of competing risks is not only complex but also difficult to interpret due to the presence of multiple causes of failure (24). Tree survival models are strong tools for analyzing survival data.

The current study is an attempt to determine the five-year survival time and factors affecting it in the presence of competing risks using tree survival model with maximum within-node homogeneity among children suffering from acute lymphoblastic leukemia.

Materials and Methods

This study extracted and collected all data associated with 424 under-15 patients suffering from ALL and referring to Dr. Sheikh Hospital in Mashhad, Iran between 2007 and 2015. Afterwards, the survival

information of the patients was elicited through telephone interview, and then registered in the checklist.

Demographic information, together with laboratory information including bone marrow aspiration report, flow cytometry, cytogenetic and CBC of the patients and information on the patients' treatment were extracted via analyzing the patients' documents. Other information was collected based on the checklist.

Age, sex, treatment period, white blood cells count, platelet count, LDH level, patients' cell line, CNS involvement, as well as clinical symptoms like mediastan mass, rheumatologic symptoms, Tumor Lysis Syndrome, Hepatosplenomegaly greater than 4 cm and Lymphadenopathy greater than 2 cm serve as the independent variables of the study.

The patients' survival time from diagnosis to the first relapse, relapse-free death, the patient's withdrawal from study or completion of follow-up period were considered on a monthly basis, and those patients who were eligible for a 60-month follow-up were included in the study. Living people at the end of study and those missed in the follow-up were considered right censor. The relapse of the disease and relapse-free death were adopted as the desired event and competing event, respectively.

Statistical Analysis

The median, mean and survival diagram were computed and depicted using Kaplan-Meier method. The factors contributing to the patients' survival were identified using tree survival model in the presence of competing risks based on maximum within-node homogeneity and using event-specific

martingale residual resulting from the fitting of Cox Proportional Hazards Model. Cumulative incidence curve was calculated for each sub-group, and the mean survival time to the relapse of the disease was compared among three sub-groups using Gray's test. Adopting a 0.05 level of significance, the current study evaluates the collected data using R 3.3.1 software together with foreign, survival, cmprsk, rpart and partykit software packages.

Results

255 patients, consisting of 147 boys (57.6%) with the average age of 5.9 years old and 108 (42.4) girls with the average age of 5.77, were eligible for the 60-month follow-up program. Among all patients, a number of 26 children (10%) experienced relapse and 57 ones (22.3%) experienced relapse-free death. 172 children (67.4%) were considered right censored. The probability of patients' survival was calculated 0.662. The truncated mean of the patients' survival time (the time expected for the lack of the occurrence of relapse and death after diagnosis) was determined 44.20 with the standard deviation of 1.52. The truncated means of the patients' post-diagnosis survival time were obtained 55.51 with SD of 0.906, and 47.53 with SD of 1.47 months, respectively. Having compared the Kaplan-Meier diagrams for relapse and relapse-free events in terms of a unique event, or two separate events (relapse or death, or both), one can come to this conclusion that the patients' survival time is lower in the presence of competing risks than in their absence (diagrams 1&2).

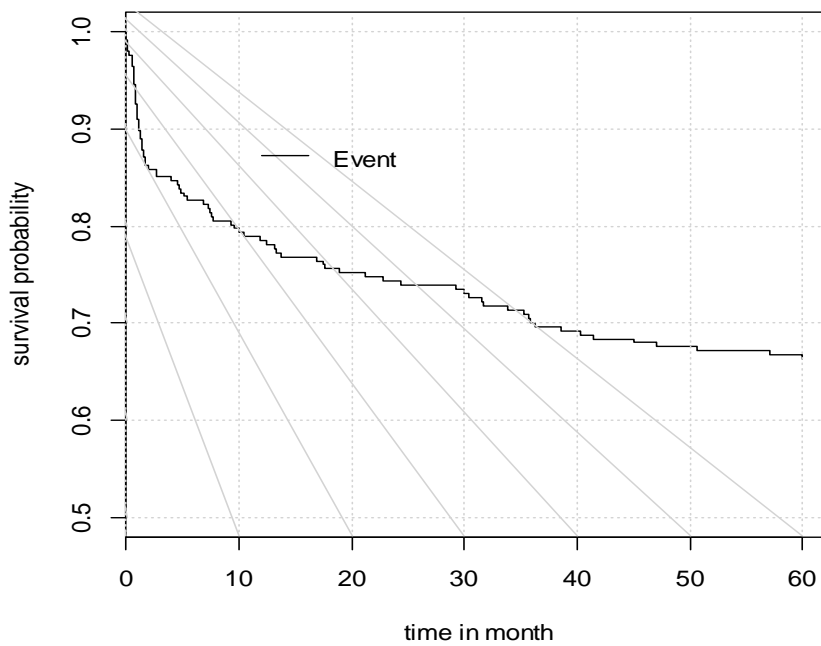


Diagram 1: Kaplan-Meier survival curve for the occurrence of death or relapse-free death in the patients with acute lymphoblastic leukemia

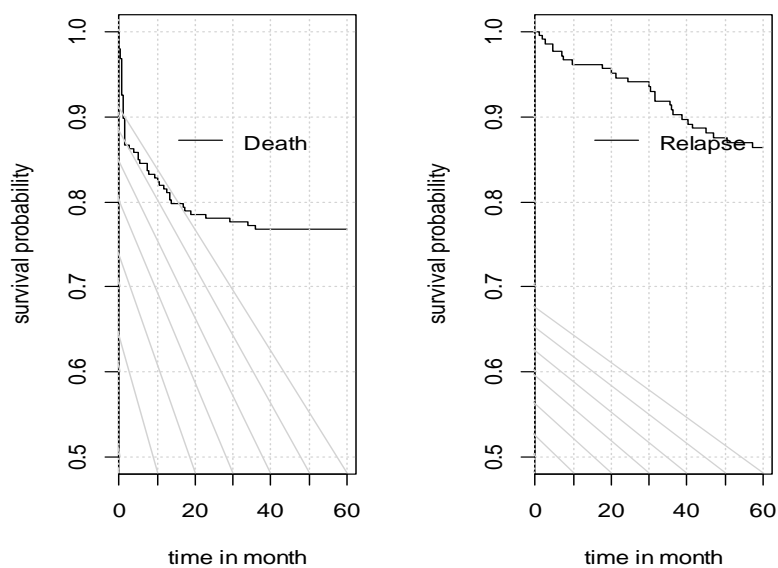


Diagram 2: Kaplan-Meier survival curve for the occurrence of death or relapse-free death in the patients with acute lymphoblastic leukemia

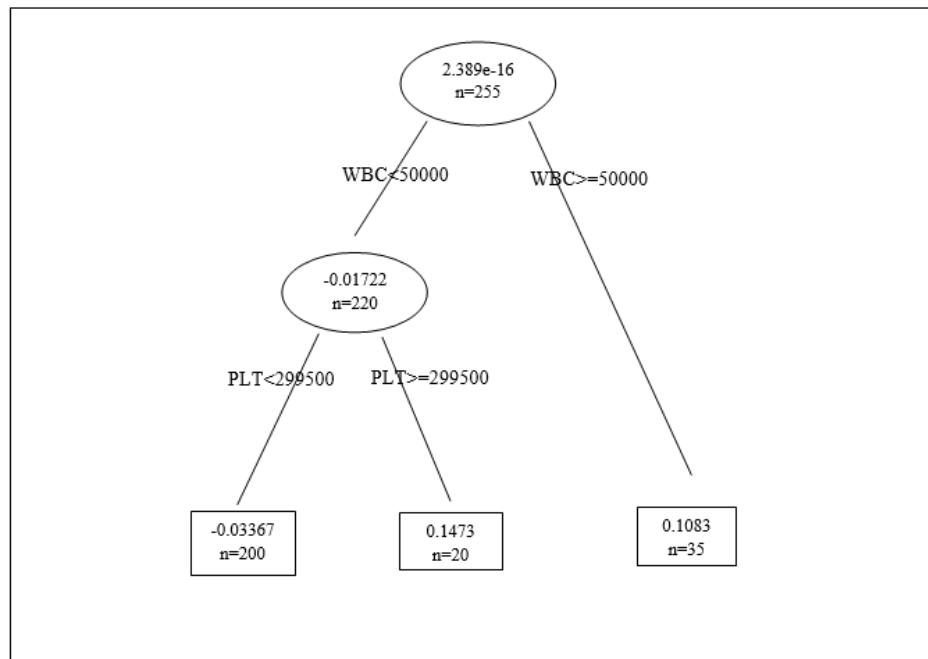


Diagram 3: Identification of the homogenous sub-groups of the patients using competing risks tree model

According to the impurity function of sum of squares of Event-Specific Martingale Residual, the patients were classified into two homogenous sub-groups based on white blood cells count (>50000 or ≤ 50000). To put it differently, regarding this variable, the first partition revealed the most improvement in the fitting of the model. The patients with WBC count lower than 50000 were divided into two homogenous sub-groups (200 patients with platelet count <299500 and 20 patients with platelet count ≥ 299500). This partition reduces data impurity in node 2, thereby increasing the homogeneity by 0.0303. The decreased deviance of this node (19.67) in comparison to the root node (25.85) implies an improvement in the goodness-of-fit of the model.

In the terminal nodes 3, 4 and 5, no partition has been implemented to create homogenous sub-groups. As such, these sub-groups of the patients constitute three separate homogenous sub-groups with respect to the duration of

disease or its relapse. The significance decline in the deviance of these nodes comparing to the root node alludes to the enhancement of the goodness-of-fit of the model. Diagram 3 represents competing risks tree diagram and patients' partitioned sub-groups together with prognostic factors of the disease.

The information of each sub-group including the number of the patients and competing risks together with the estimated survival time is illustrated in Table 2. The patients with WBC count lower than 50000 and platelet count lower than 299500 were proved to have longer survival time than the other sub-groups. Moreover, the cumulative incidence curve of recurrent events was estimated for each sub-group, and then the mean survival time to the relapse of the disease was compared among three sub-groups. The comparison made among the cumulative incidence curves indicates that the mean survival time to the relapse of the disease for the homogenous sub-groups of the patients with acute lymphoblastic leukemia is varied.

The results of Gray’s test to compare the survival time among three sub-groups point to the significance discrepancies among these

three sub-groups regarding the mean survival time to the relapse of the disease ($p=0.0055$).

Table 1: The information of each node based on the fitting of the terminal competing risks tree.

Name	Partition variable	Observations	deviance statistics	The mean of fitted model residuals	The mean square of errors	The Improvement rate
1	WBC	255	25.85	$2.389e^{-16}$	0.101	0.0184
2	PLT	220	19.67	-0.0172	0.089	0.0303
3	Terminal node	200	14.89	-0.0337	0.075	-
4	Terminal node	20	4.18	0.1473	0.209	-
5	Terminal node	35	5.71	0.1083	0.163	-

Table 2: The information of the homogenous sub-groups based on the fitting of the competing risks tree model.

Sub-group (terminal node)	Observations (percent)				The Estimated Mean survival rate (month)	The standard error estimate
	sum	Relapse	Relapse-free death	Censored		
1(Node3) WBC < 50000 & LT < 29950	200	14(7)	46(23)	140(70)	42.16	25.22
2(Node4) WBC < 50000 PLT ≥ 29950	20	5(25)	3(15)	12(60)	41.54	22.66
3(Node5) WBC ≥ 50000	35	7(20)	8(22.9)	20(57.1)	36.83	26.41

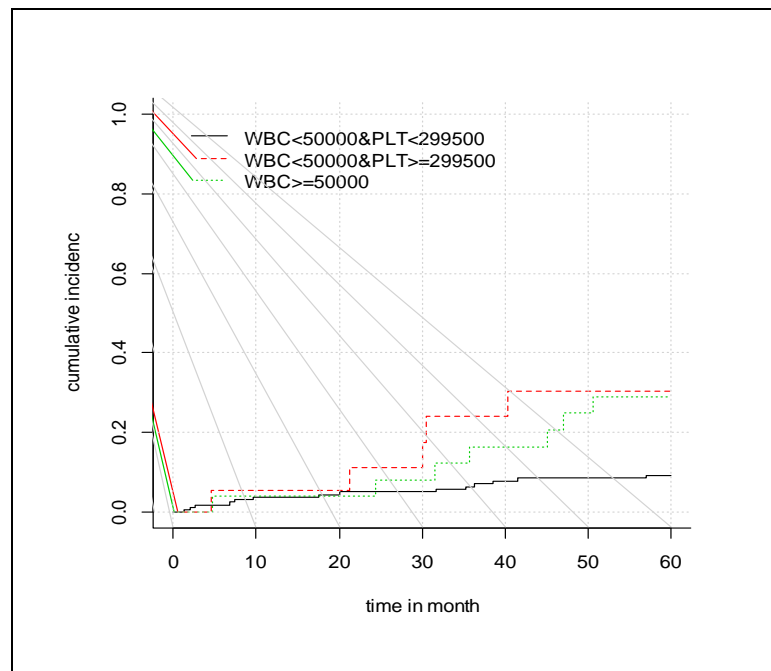


Diagram 4: The cumulative incidence curve of recurrent events among homogenous sub-groups of the patients with acute lymphoblastic leukemia

Discussion

The present study was concerned with exploring the WBC and platelet count as important factors contributing to the survival time to the relapse of the disease. Zareifar et al employed Cox regression model and reported platelet and the number and history of relapses as major factors influencing the survival of the under-15 children who suffered from acute lymphoblastic leukemia and referred to Shahid Faghihi Hospital, in Shiraz, Iran (26).

Having employed Cox and Weibull proportional hazards models, Hosseinitashnizi et al made an attempt to explore the factors affecting the survival time of patients with acute lymphoblastic leukemia and concluded that variables such as age, hometown and WBC count exert significant impact on the patients' survival time in either models. The

distinction between these two models can be associated with the patients' age and the effect of certain variables like hometown (27). Chessells et al used Kaplan-Meier method and ran Long-rank test on 378 patients suffering from acute lymphoblastic leukemia and accentuated age and WBC as factors influencing patients' survival (28). Hazar and his colleagues conducted a study on 142 Turkish children aged 1 to 18 years old using Cox regression model and documented that age ≤ 10 years old, WBC count, Hepatosplenomegaly larger than 4 cm, medistian mass, and patients' cell line are amongst leading factors causing mortality. Among identified factors, the only factor that overlaps with our study is WBC count. The way of classifying the patients based on their age seems to be the cause of distinction in the results of the studies (29). In a study undertaken by Hashemi et al using Cox

proportional hazards model, $WBC \geq 50000$ was introduced as a major factor influencing the survival time of the patients with acute lymphoblastic leukemia referring to Shahid Sadoghi hospital in Yazd, Iran. Their findings conform to the results of the current study (30). Chen et al investigated 78 Taiwanese children suffering from acute lymphoblastic leukemia using cox proportional hazards model and pointed to $WBC \geq 50000$ as a major contributor to the survival time of the patients (31). To detect the factors influencing the patients' survival time, Mousavinasab et al carried out a study on 97 children with acute lymphoblastic leukemia in Boualisina hospital, in Sari, Iran using competing risks regression and Cox proportional hazards models. They concluded upon the whole that the competing risks model introduces variables like leukemia type, WBC count and hemoglobin level as contributors affecting the survival time, whereas analyses based on Cox proportional hazards model reported the same impact as a result of WBC count, age and leukemia type (32). Bahrami et al examined the survival time of 52 patients died of acute lymphoblastic leukemia in Seyedoshohada hospital in Isfahan, Iran using EM algorithm and Monte Carlo simulation method. They showed a significant relationship between immature cells and hemoglobin, and patients' survival time. Contrary to their findings, the present research found out that variables like WBC and platelet count influenced the studied patients. It is noteworthy that Bahrami et al carried out their study on all age ranges, different from what being considered in the current study (33). To study factors influencing the survival time of the patients with acute lymphoblastic leukemia in Dr. Sheikh hospital, in Mashhad, Iran during the years 2007-2015 using competing risks, Bonakchi et al fitted Cox proportional hazards and subdistribution models based on

univariate and multivariate analyses. According to the Proportional Hazards Model for the Subdistribution, variables like WBC and platelet count influence the relapse of cancer in the sampled patients. Similarly, the analysis based on the classified Cox model indicated that besides WBC, hemoglobin and Tumor Lysis Syndrome, the rest of variables such as platelet count, cell line, Hepatosplenomegaly and mediastinal mass did not show the same effect on the patients' survival time, thereby confirming the results of the present study (34).

In addition, the findings of the current study recognized three sub-groups at risk with respect to the survival time. In the third sub-group, the patients with $WBC \geq 50000$ showed the least mean survival time. In fact, these models aim at establishing sub-groups of patients, whose members represent the maximum homogeneity in terms of survival time. The mean survival time of the patients in the presence of the events of relapse-free death and relapse was estimated as a unique event, hence equaling 44.20. Also, the mean survival time of relapse-free death and relapse of the disease was obtained 55.51 and 47.53, respectively. Therefore, the patients' survival time in the presence of competing events is lower than that in the absence of competing event.

Conclusion

The ongoing research employed a tree model with maximum within-node homogeneity based on the impurity function of sum of squares of Event-Specific Martingale Residual (relapse of the disease) and ranked the survival time of the patients with acute lymphoblastic leukemia recursively with respect to certain demographic, laboratory and clinical specifications. The results of the fitting of the tree model pointed to three

homogenous sub-groups based on two prognostic factors affecting the mean survival time to the relapse of the disease. Among fourteen studied factors, white blood cell count and platelet count were identified as the most influential prognostic factors in

Acknowledgements

This article was derived from a Master's thesis submitted by Somaye Barzanouni in biostatistics with the project code of 940972. This work would not have been possible without the financial and non-financial supports of the deputy of research and technology of Mashhad University of Medical Sciences, especially the faculty of hematology and oncology. We are really indebted to them. The last but not least, we are also grateful to Mrs. Zahra Payab, the nurse of the department

References

1. Alsayyad J, Hamadeh R. Cancer incidence among the Bahraini population: a five-year (1998–2002) experience. *Hematology/oncology and stem cell therapy*. 2008;1(3):175-82.
2. del Pilar Díaz M, Osella AR, Aballay LR, Muñoz SE, Lantieri MJ, Butinof M, et al. Cancer incidence pattern in Cordoba, Argentina. *European Journal of Cancer Prevention*. 2009;18(4):259-66.
3. Mousavi SM, Gouya MM, Ramazani R, Davanlou M, Hajsadeghi N, Seddighi Z. Cancer incidence and mortality in Iran. *Annals of oncology*. 2008;20(3):556.63-6.(persian)
4. Bao PP, Zheng Y, Wang CF, Gu K, Jin F, Lu W. Time trends and characteristics of childhood cancer among children age 0–14 in Shanghai. *Pediatric blood & cancer*. 2009;53(1):13-6.
5. Buka I, Koranteng S, Vargas ARO. Trends in childhood cancer incidence: review of environmental linkages. *Pediatric Clinics of North America*. 2007;54(1):177-203.
6. Michel G, Von Der Weid N, Zwahlen M, Redmond S, Strippoli MP, Kuehni C. Incidence of

predicting the patients' survival time. Therefore, these controllable factors are recommended to be considered in patients' therapeutic protocols to prolong the survival time of the patients with acute lymphoblastic leukemia.

of oncology in Taleghani Hospital, Gorgan, Iran.

Declarations

Conflict of interest

We declare that we have no financial or non-financial conflicts of interest related to the subject matter or materials discussed in the article.

Authors' contributions

All authors contributed equally to this work.

childhood cancer in Switzerland: The Swiss childhood cancer registry. *Pediatric blood & cancer*. 2008;50(1):46-51.

7. Moradi A, Semnani S, Roshandel G, Mirbehbehani N, Keshtkar A, Aarabi M, et al. Incidence of childhood cancers in golesstan province of Iran. *Iranian journal of pediatrics*. 2010;20(3):335.(persian)

8. Stack M, Walsh PM, Comber H, Ryan CA, O'Lorcain P. Childhood cancer in Ireland: a population-based study. *Archives of disease in childhood*. 2007.

9. Yang L, Fujimoto J, Qiu D, Sakamoto N. Childhood cancer in Japan: focusing on trend in mortality from 1970 to 2006. *Annals of oncology*. 2008;20(1):166-74.

10. Leukemia CL. The leukemia and lymphoma society. In *TerneT SITE*. 2008.

11. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Murray T, et al. *Cancer statistics, 2008*. CA: a cancer journal for clinicians. 2008;58(2):71-96.

12. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. *CA: a cancer journal for clinicians*. 2009;59(4):225-49.
13. Jemal A, Siegel R, Ward E, Murray T, Xu J, Smigal C, et al. Cancer statistics, 2006. *CA: a cancer journal for clinicians*. 2006;56(2):106-30.
14. Jemal A, Siegel R, Ward E, Murray T, Xu J, Thun MJ. Cancer statistics, 2007. *CA: a cancer journal for clinicians*. 2007;57(1):43-66.
15. Karimi M, Mehrabani D, Yarmohammadi H, Jahromi FS. The prevalence of signs and symptoms of childhood leukemia and lymphoma in Fars Province, Southern Iran. *Cancer detection and prevention*. 2008;32(2):178-83(persian).
16. Coebergh J-W, Pastore G, Gatta G, Corazziari I, Kamps W, Group EW. Variation in survival of European children with acute lymphoblastic leukaemia, diagnosed in 1978–1992: the EUROCARE study. *European Journal of Cancer*. 2001;37(6):687-94.
17. Sajjadi H, Roshanfekr P, Asangari B, Gharai N, Torabi F. Quality of life and satisfaction with services in caregivers of children with cancer. *Iran Journal of Nursing*. 2011;24(72):8-17.(persian)
18. Malempati S, Gaynon PS, Sather H, La MK, Stork LC. Outcome after relapse among children with standard-risk acute lymphoblastic leukemia: Children's Oncology Group study CCG-1952. *Journal of Clinical Oncology*. 2007;25(36):5800-7.
19. Arellano ML, Langston A, Winton E, Flowers CR, Waller EK. Treatment of relapsed acute leukemia after allogeneic transplantation: a single center experience. *Biology of blood and marrow transplantation*. 2007;13(1):116-23.
20. Lawson S, Harrison G, Richards S, Oakhill A, Stevens R, Eden O, et al. The UK experience in treating relapsed childhood acute lymphoblastic leukaemia: a report on the medical research council UKALLR1 study. *British journal of haematology*. 2000;108(3):531-43.
21. Kazak AE, Barakat LP, Meeske K, Christakis D, Meadows AT, Casey R, et al. Posttraumatic stress, family functioning, and social support in survivors of childhood leukemia and their mothers and fathers. *Journal of consulting and clinical psychology*. 1997;65(1):120-9.
22. Klein JP, Zhang M-J. Survival analysis. *Handbook of Statistics*. 2007;27:281-317.
23. Abadi A, Dehghani-Arani M, Yavari P, Alavi-Majd H, Bajik K. Application of the competing risk models for the analysis of risk factors in patients with breast cancer. *Feyz Journals of Kashan University of Medical Sciences*. 2013;16(6).(persian)
24. Saki Malehi A, Hajizadeh E, Fatemi R. Evaluation of prognostic variables for classifying the survival in colorectal patients using the decision tree. *Iranian Journal of Epidemiology*. 2012;8(2):13-9.(persian)
25. Bacchetti P, Segal MR. Survival trees with time-dependent covariates: application to estimating changes in the incubation period of AIDS. *Lifetime data analysis*. 1995;1(1):35-47.
26. Zarei far S, Hashiani almasi A, Karimi M, Tabatabai S, Qyasvnd R. The five-year survival rate and its influencing factors in pediatric leukemia. *Koomesh*. 1391;14(1).
27. Hosseini Teshnizi S, Zare S, Tazhibi M. The evaluation of Cox and Weibull proportional hazards models and their applications to identify factors influencing survival time in acute leukemia. In: *Sci JHUM*, editor. 2012. p. 269-78.(persian)
28. Chessells J, Hardisty R, Richards S. Long survival in childhood lymphoblastic leukaemia. *British journal of cancer*. 1987;55(3):315-9.
29. Hazar V, Karasu GT, Uygun V, Akcan M, Küpesiz A, Yesilipek A. Childhood acute lymphoblastic leukemia in Turkey: factors influencing treatment and outcome: a single center experience. *Journal of pediatric hematology/oncology*. 2010;32(8):e317-e22.
30. Hashemi AS, Manuchehri Naini MA, Islami Z, Lotfi MH, Khairandish M, Rafieean M. Immunofenotype in pediatric patients with acute lymphoblastic leukemia referring to Shahid Sadoughi Hospital in YazdAzam Sadat Hashemi. *Journal of Shahid Sadoughi University of Medical Sciences and Health Services*. 2007;16(5):56-60.(persian)
31. Chen B-W, Lin D-T, Lin K, Chuu W, Su S, Lin K. An analysis of risk factor and survival in childhood acute lymphoblastic leukemia. *Zhonghua Minguo xiao er ke yi xue hui za zhi*

[Journal] Zhonghua Minguo xiao er ke yi xue hui . 1989;30(5):299-308

32.Mousavinasab SN, Yazdani Cherati J, Karami H, Khaksar S. Risk Factors Influencing the Survival of Pediatric Acute Leukemia Using Competing Risk Model. Journal of Mazandaran University of Medical Sciences. 2015;24(121):31-8.(persian)

33.Bahrami M, Moshkani MR, Alam Samimi M. Evaluation of effective factors on the survival

time of patients with acute leukemia and the estimated mean survival time by Expectation & Maximization algorithm and Monte Carlo Markov simulation method. Journal of Isfahan Medical School. 2006;25(84).(persian)

34.Bonakchi H, Farhangi H, Esmaily H, Boosti H, Forouzannejhad M. Factors Affecting Survival of Children with Acute Lymphoblastic Leukemia Using Competing Risks Model. ZUMS Journal. 2017;25(110):123-36.(persian)

How to cite:

Esmaily H, Barzanouni S, Farhangi H. Determining factors contributing to the five-year survival of children suffering from acute lymphoblastic leukemia based on tree survival model in the presence of competing risks. Jorjani Biomedicine Journal. 2018; 6(4): 29-39.