Original Article
http://mjiri.iums.ac.ir
Medical Journal of the Islamic Republic of Iran (MJIRI)

Med J Islam Repub Iran. 2017(17 Dec);31.97. https://doi.org/10.14196/mjiri.31.97

Survival analysis of thalassemia major patients using Cox, Gompertz proportional hazard and Weibull accelerated failure time models

Enayatollah Bakhshi[1], Reza Ali Akbari Khoei[1], Azita Azarkeivan[2], Maryam Kooshesh[3], Akbar Biglarian[4]*

Received: 17 May 2017 Published: 17 Dec 2017

Abstract

Background: Thalassemia major (TM) is a severe disease and the most common anemia worldwide. The survival time of the disease and its risk factors are of importance for physicians. The present study was conducted to apply the semi-parametric Cox PH model and use parametric proportional hazards (PH) and accelerated failure time (AFT) models to identify the risk factors related to survival of TM patients.

Methods: The data of this historical cohort study (296 patients with TM) were collected during 1994 and 2013 in Zafar Clinic in Tehran. Gompertz PH and Weibull AFT models were used for survival analysis (SA) of these patients. Data analysis was performed using R3.2.2 software.

Results: 153 (51.7%) of patients were female; the mean (±SD) age of the patients was 29.11 (±0.47) years. One-year survival rate for males and females was 0.963±0.007 and 0.973±0.013, respectively; and 3-year survival rate for males and females was 0.711±0.057 and 0.733±0.114, respectively. In the Gompertz model, birthplace and age at onset of the disease were significant factors (p= 0.035, and p= 0.005) in survival time. Also, in the Weibull model, birth place and age at onset of the disease were significant factors (p= 0.013, and p= 0.008) in survival time. The Akaike Information Criterion (AIC) for Weibull model was 158.51, which was lower than other parametric models.

Conclusion: According to the results, the Weibull AFT model was found to be a better model for identifying the risk factors related to survival of patients with TM disease. Informing parents, especially mothers and paying attention to blood screening for early diagnosis may increase the survival rate of patients.

Keywords: Survival analysis, Accelerated failure time model, Proportional hazards models, Thalassemia major

Copyright© Iran University of Medical Sciences

Cite this article as: Bakhshi E, Ali Akbari Khoei R, Azarkeivan A, Kooshesh M, Biglarian A. Survival analysis of thalassemia major patients using Cox, Gompertz proportional hazard and Weibull accelerated failure time models. *Med J Islam Repub Iran. 2017* (17 Dec);31:97. https://doi.org/10.14196/mjiri.31.97

Introduction

Thalassemia is the most common inherited anemia in the world and also in Iran (1, 2). Middle Eastern countries have a high concentration of thalassemia (2). Moreover, 18 000 deaths occurred due to thalassemia major (TM) in 2010 (3). It has been reported that 5.2% of the world population are carriers of such disorders (4) and also the prevalence rate of thalassemia carriers is 4.5% in Iran (5). In addition, it has been reported that approximately 300 children with TM are born in Iran and also over 264 000 affected infants are born annually in developing countries

Corresponding author: Dr Akbar Biglarian, abiglarian@uswr.ac.ir

1. Department of Biostatistics, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran.
2. Pediatric Hematology Oncology, Iranian Blood Transfusion Organization (IBTO) – High Institute for Research and Education in Transfusion Medicine, Thalassemia Clinic, Tehran, Iran.
3. Tehran University of Medical Sciences, Tehran, Iran.
4. Department of Biostatistics, Social Determinants of Health Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran.

↑What is "already known" in this topic:
Most data analysts have used the Cox PH model, as the most and popular semi-parametric methods, to analyze the survival of TM patients. This model does not make specific assumptions about the probability distribution of event times.

→What this article adds:
The Weibull AFT model is a more general parametric method that can be used for survival analysis of thalassemia major (TM patients). This model has not been used in this field in Iran. However, based on this model, informing parents, especially mothers, and paying attention to blood screening for early diagnosis were most important related factors in the survival rate of TM patients.

[Downloaded from mjiri.iums.ac.ir on 2024-03-29]

[DOI: 10.14196/mjiri.31.97]

(4-6). The thalassemia program was implemented in 1997 in Iran and aimed at reducing and preventing this disease (5, 6). Identifying the risk factors that affect the survival rate of TM patients is highly important for physicians and healthcare systems.

In this regard, survival analysis (SA) can be used to identify the risk factors for survival of patients (2, 6). This analysis, ie, SA, is a collection of statistical methods that involves the modeling of time to event data and is used to predict the occurrence and timing of events (7). In SA, which typically focuses on time to event data, there are different models to analyze survival times of data based on a set of predictor variables (7, 8), such as proportional hazard (PH) and accelerated failure time (AFT) models (7, 9). In the previous decades, data analysts have used a various survival methods to determine prognostic factors and survival rate of TM patients (1, 2, 6, 10-16). The Cox PH model is one of the most popular semi-parametric methods used to predict survival of TM patients (7). The Weibull AFT model is a parametric method which can be used to predict the survival of TM patients. This model is more general than other parametric models (7, 8).

The most popular method for regression analysis of survival data is Cox regression, which combines the proportional hazards model with the partial likelihood method of estimation. It does not make specific assumptions about the probability distribution of event times. By contrast, parametric regression models assume particular families of probability distributions, such as exponential, Weibull, Gompertz, lognormal, log-logistic, or gamma (7, 9, 17).

However, it is of paramount importance to identify the prognostic factors affecting the long-term survival rate of TM patients. This study aimed at applying semi-parametric Cox PH model and a parametric PH and AFT model to identify the risk factors of TM patients.

Methods

Patient population

This historical cohort study was conducted by enrolling 296 TM patients. These patients had been referred to Zafar Thalassemia Clinic in Tehran during 1994 and 2013. This study was approved by the ethical committee of University of Social Welfare and Rehabilitation Sciences (IR.USWR.REC.1395.5).

Data collection

Data of all patients were collected based on patients' medical records, and the collected data included the following information: sex, birth place (near the sea/not), date of birth (after 1997/before 1997), age at onset of disease, age of Desferal injection, age of transfused blood, splenectomy (no/yes), maternal education (no/yes), and consanguinity (no/yes). The survival time of each patient, in year, was defined as the interval between the date of birth and the date of death or date of follow- up time in May 2013 (the end of the study).

Statistical analyses

Cumulative survival rates were estimated using the Kaplan-Meier method. The log-rank test was used to compare the survival curves across the subgroups. A multivariate (adjusted analysis) Cox proportional hazards model was utilized to identify risk factors relating to survival of patients. Finally, parametric Gompertz PH and Weibull AFT model were used to determine the SA of TM patients. The Akaike information criterion (AIC) was utilized to compare these parametric models. The smaller value for this criterion means a better and more suitable model (7). All statistical analyses were performed using R 3.2.2 software. A $p<0.05$ was considered statistically significant.

AFT model

AFT model is a parametric model that assumes that the effect of a covariate is to accelerate or decelerate the survival (or the hazard) of a disease over time by some constant. For i = 1, …, n, let T_i, C_i and X_i would be the failure time, censoring time, and the p×1 covariate vector for the i-th subject. It is assumed that T_i is conditionally independent of C_i given X_i. An univariate semiparametric AFT model has the form $\log(T_i) = X_i^T \beta + \varepsilon_i \quad i = 1,...,n,$ where β is an unkhnown p×1vector of regression parameter, ε_i's are independent of X_i and also independent, which is identically distributed in random variables with an unspecified distribution. Imagine that in i-th cluster, there is a random sample of n independent cluster with K_i margin. In this case, the multivariate AFT model has the following form: $\log(T_{ik}) = X_{ik}^T \beta + \varepsilon_{ik} \quad i = 1,...,n, k = 1,...,K_i$. In the presence of censoring, the observed data would consist of copies of $\left(Y_{ik}, \Delta_{ik}, X_{ik}^T\right)$, where $Y_{ik} = \min(T_{ik}, C_{ik})$ and $\Delta_{ik} = I(T_{ik} < C_{ik})$, and I(.) is the indicator function. Without censoring, the ordinary least squre (OLS) method can be used to estimate all parameters. With censoring, one can use the Buckly-James estimator by adapting the OLS method (18-20).

Results

A total of 296 TM patients were included, with total time at risk, 2105.5 years. The minimum and maximum follow-up times were 5 and 49 months, respectively. Overall, 38 (12.84%) patients died. A total of 153 patients were female; and the mean±SD of patients' age was 29.11±0.47 years. Overall, 2 patients had irregular blood transfusions, 141 (47.4%) had near the sea birth place, 215 (72.6%) were born after 1997, 112 (37.8%) had splenectomy, and consanguinity was observed in 62 (21.6%) patients. In addition, 46 (15.6%) parents were illiterate. The mean (and median) ± SD of age at onset of disease, age of

Table 1. One-, two-, and three- year survival probability of patients with TM using Kaplan-Meier method

Years	Total	Women	Men
One-year	0.969±0.010	0.973±0.013	0.963±0.007
Two-year	0.882±0.021	0.902±0.028	0.856±0.033
Three-year	0.728±0.046	0.733±0.114	0.711±0.057

[Downloaded from mjiri.iums.ac.ir on 2024-03-29]

[DOI: 10.14196/mjiri.31.97]

Table 2. Results of the Cox PH model for TM patients

Variables	Parameter estimation	SE*	P-value	HR
Birth place (near the sea)	-1.642	0.897	0.067	0.193
Age at onset of disease	0.320	0.156	0.040**	1.377
Age of Desferal injection	-0.027	0.109	0.807	0.974
Splenectomy (no)	-0.320	0.775	0.679	0.726
Maternal education	-0.431	0.470	0.358	0.650
Consanguinity (yes)	0.693	0.698	0.321	2.000
Age of transfused blood	-0.434	0.241	0.075	0.648
Born time (after 1997)	1.402	1.190	0.238	4.062
Gender (female)	-0.994	0.700	0.156	0.370
AIC=114.460	-	-	-	-

*Standard error; **Significant at alpha=0.05

Table 3. Results of the Gompertz PH model for TM patients

Variables	Parameter estimation	SE*	P-value	HR
Birth place (near the sea)	-1.860	0.884	0.035**	0.156
Age at onset of disease	0.530	0.190	0.005**	1.699
Age of Desferal injection	-0.074	0.098	0.450	0.928
Splenectomy (no)	-0.375	0.757	0.620	0.687
Maternal education	-0.426	0.446	0.339	0.653
Consanguinity (yes)	0.816	0.684	0.233	2.260
Age of transfused blood	-0.435	0.242	0.073	0.647
Born time (after 1997)	1.760	1.140	0.123	5.810
Gender (female)	-0.994	0.681	0.144	0.370
Shape parameter	0.071	0.037	-	-
Rate(or scale) parameter	0.003	0.004	-	-
AIC=161.769	-	-	-	-

*Standard error; **Significant at alpha=0.05

Table 4. Results of the Weibull AFT model for TM patients

Variables	Parameter estimation	SE*	P-value	HR
Birth place (near the sea)	0.243	0.098	0.013**	1.275
Age at onset of disease	-0.246	0.092	0.008**	0.782
Age of Desferal injection	0.036	0.044	0.422	1.036
Splenectomy (no)	0.172	0.351	0.625	1.187
Maternal education	0.195	0.209	0.350	1.216
Consanguinity (yes)	-0.384	0.324	0.236	0.681
Age of transfused blood	0.202	0.117	0.084	1.223
Born time (after 1997)	-0.776	0.534	0.146	0.460
Gender (female)	0.464	0.325	0.154	1.590
Shape parameter	3.976	0.722	-	-
Scale parameter	0.781	0.181	-	-
AIC=158.509	-	-	-	-

*Standard error; **Significant at alpha=0.05

Desferal injection and age of transfused blood was 2.31(0.83)±4.90, 5.79(4.00)±5.94, and 1.72(1.00)±2.82, respectively.

Three-year survival rate for males and females was 0.711±0.057 and 0.733±0.114, respectively, based on Kaplan-Meier analysis (Table 1). The mean ± SD survival time of males and females was 41.75±1.46 and 37.43 ±0.76, respectively. The log-rank test revealed no difference between the mean survival time of males and females ($X^2 = 3.197$, df= 1, p= 0.074).

In the Cox proportional hazard model, age at onset of disease was a significant factor (p=0.040) in survival time (Table 2). In the Gompertz model, variables of birth place and living near the sea were significant factors (p= 0.035 and p = 0.005, respectively) in survival time (Table 3). In the Weibull model, birthplace and age at onset of disease were significant variables (p= 0.013 and p= 0.008, respectively) in survival time (Table 4). The AIC for Weibull model was 158.51, which was lower than the Gompertz model (161.8). For this model, the hazard ratio (HR) estimate for age at onset of disease was exp (-0.246) = 0.782. Adjusted for other variables, this means that the age at onset of disease has a hazard rate of about 0.78 times on the hazard of the event, if 1 unit increased in the age at onset of disease. The hazard ratio estimate for birth place was exp (0.243)= 1.28, meaning that near the sea place of birth has a hazard rate of about 1.3 times more than birth at other places on the hazard of the event.

Discussion

A total of 81.5% of the TM patients survived in the study period, which is the overall median survival time that could not be calculated. However, the 1-year, 2-year, and 3-year survival rates of the patients were 97%, 88%, and 78%, respectively. On the other hand, TM patients had a high survival rate, and improvement in the survival rate of these patients was palpable. Previous studies in Iran showed that only 50% to 60% of beta-TM patients had long-term survival rate (2, 6, 10, 15).

SA of thalassemia patients was performed using parametric and Cox models. A survival study revealed that the survival rate of beta-TM patients was affected by accompanied diseases, kind of transfused blood, and rural residency (2). The 20-year- survival rate was reported to be

[Downloaded from mjiri.iums.ac.ir on 2024-03-29]

[DOI: 10.14196/mjiri.31.97]

85%; moreover, education, marital status, ferritin levels, and comorbidities factors were related to survival rate of thalassemia patients (6). In another study, regular blood transfusion, blood screening, and family awareness affected the survival rate of these patients (10). The survival of those patients who were born after 1981 was better than the others (11). The 20-year- survival rate of TM was reported to be 98.5%. Multivariate survival analysis revealed that female sex significantly affected the survival time, but not the genotype, birth cohort (after 1974), or treatment center (12). Death of TM patients was reported to be 50% before the age of 35. Moreover, improved survival rate and survival without complication due to treatment was also reported (13, 14). The survival of TM patients was investigated and it was found that ethnicity and gender did not have any significant effects on the survival of these patients (1). Being alive until 20 years old was reported in 68% of TM patients and only half of them lived until 30 years old (15). The Cox regression model revealed that the interaction effect of sex and age was a risk factor for predicting mortality in patients (18). The expected survival probability of patients at 50 years was reported to be 0.63 (19). A study has reported better median survival time (4.89 yrs. for females and 3.84 yrs. for males) and better prognosis for females (20). The survival of both genders for birth cohort before and after 1986 was reported without any statistically significant differences. The survival rate of patients who were born after 1986 was reported to be better at 30 years old (68% vs. 80%) (21).

Conclusion

This study revealed that the variables of birth place, age at onset of disease, and age of transfused blood had significance effects on the survival of these patients. Unlike other studies, birth time of patients and gender did not significantly affect the survival of patients. One of the limitations of this study was missing values in explanatory variables. Based on the results, Weibull parametric model was selected as the best model. So, this model is suggested to other researchers. Furthermore, informing parents, especially mothers, and paying attention to blood screening for early diagnosis may increase the survival rate of patients.

Acknowledgements

We wish to express our special thanks to all colleagues at Zafar Clinic in Tehran for data collection. Also, we thank the Deputy of Research and Technology of University of Social Welfare and Rehabilitation Sciences for financial support.

Conflict of Interests

The authors declare that they have no competing interests.

References

1. Latifi S, Zandian K. Survival analysis ofβ-thalassemia major patients in Khouzestan province referring to Shafa hospital [in Persian]. Jundishapur Sci Med J. 2010;9(1):84-92
2. Zamani R, Khazaei S, Rezaeian S. Survival analysis and its associated factors of Beta thalassemia major in Hamadan province. Iran J Med Sci. 2015;40(3):233-239.
3. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012;380:2095-2128.
4. Samavat A, Modell B. Iranian national thalassaemia screening programme. BMJ. 2004;329:1134-1137.
5. Dehshal MH, Ahmadvand A, Darestani SY, Manshadi M, Abolghasemi H. Secular trends in the national and provincial births of new thalassemia cases in Iran from 2001 to 2006. Hemoglobin. 2013; 37:124-137.
6. Rajaeefard A, Hajipour M, Tabatabaee HR, Hassanzadeh J, Rezaeian S, Moradi Z, and et al. Analysis of survival data in thalassemia patients in Shiraz, Iran. Epidemiol Health. 2015;37:e2015031.
7. Kleibaum David G, Klein M. Survival analysis: a self-learning text. 3rd ed. NewYork: Springer; 2012.
8. Patel K, Kay R, Rowell L. Comparing proportional hazards and accelerated failure time models: an application in influenza. Pharm Stat. 2006;5(3):213-24.
9. RosthøjS, Keiding N. Explained variation and predictive accuracy in general parametric statistical models: the role of model misspecification. Probability, Statistics and Modelling in Public Health: Springer; 2006. p. 392-404.
10. Roudbari M, Soltani-Rad M, Roudbari S. The survival analysis of beta thalassemia major patients in South East of Iran. Saudi Med J. 2008;29:1031-1035.
11. Ladis V, Chouliaras G, Berdousi H, Kanavakis E, Kattamis C. Longitudinal study of survival and causes of death in patients with thalassemia major in Greece. Annals New York Acad Sci. 2005;1054:445-450.
12. Telfer P, Coen Pg, Christou S, Hadjigavriel M, Kolnakou A, Pangalou E, et al. Survival of medically treated thalassemia patients in Cyprus. Trends and risk factors over the period 1980-2004. Haematologica. 2006;91(9):1187-1192.
13. Borgna-Pignatti C, Cappellini MD, De Stefano P, Del Vecchio GC, Forni GL, Gamberini MR, et al. Survival and complications in thalassemia. Annals New York Acad Sci. 2005;1054:40-47.
14. Borgna-Pignatti C, Rugolotto S, De Stefano P, Zhao H, Cappellini MD, Del Vecchio GC, et al. Survival and complications in patients with thalassemia major treated with transfusion and deferoxamine. Haematologica. 2004;89(10):1187-1193.
15. Yavarian M, Farsheedfar GR, Karimi M, Almoazzez M, Harteveld CL, Giordano PC. Survival analysis of transfusion dependent β-thalassemia major patients. J Res Health Sci. 2006;6(1): 8-13.
16. Gohari MR, Biglarian A, Bakhshi E, Pourhoseingholi MA. Use of an artificial neural network to determine prognostic factors in colorectal cancer patients. Asian Pac J Cancer Prev. 2011;12(6):1469-1472.
17. Mashhadi MA, Heidari Z, Sepehri Z, Bakhshipour AR, Karimkoshte A. The selenium status in thalassemia patients in South East of Iran. Int J Hematol Oncol Stem Cell Res. 2014;8(4):1-4.
18. Hsu CH, Taylor JM, Hu C. Analysis of accelerated failure time data with dependent censoring using auxiliary variables via nonparametric multiple imputation. Stat Med. 2015;34(19):2768-80.
19. Chen YQ. Accelerated hazards regression model and its adequacy for censored survival data. Biometrics. 2001;57(3):853-60.
20. Chiou SH, Kang S, Yan J. Fitting Accelerated Failure Time Models in Routine Survival Analysis with R Package aftgee. J Stat Software. 2014;61(11):1-23.
21. Maggio A, Vitrano A, Capra M, Cuccia L, Gagliardotto F, Filosa A, et al. Improving survival with deferiprone treatment in patients with thalassemia major: a prospective multicenter randomised clinical trial under the auspices of the Italian Society for Thalassemia and Hemoglobinopathies. Blood Cells Mol Dis. 2009;42(3):247-251.
22. Weidlich D, Kefalas P, Guest JF. Healthcare costs and outcomes of managing β-thalassemia major over 50 years in the United Kingdom. Transfusion. 2016;56(5):1038-45.
23. Abdul Nasir J, Zaidi SA. Modelling survival data of thalassaemia patients in Pakistan. J Ayub Med Coll Abbottabad. 2009;21(1):142-145.
24. Kosaryan M, Vahidshahi K, Karami H, Forootan MA, Ahangari M. Survival of thalassemic patients referred to the Boo Ali Sina Teaching Hospital, Sari, Iran. Hemoglobin. 2007;31(4):453-462.

[Downloaded from mjiri.iums.ac.ir on 2024-03-29]

[DOI: 10.14196/mjiri.31.97]