



Research article

Parametric and the Cox risk model in the analysis of factors affecting the time of diagnosis of retinopathy with patients type 2 diabetes

Fatemeh keshavarzi¹, Mohsen Askarishahi^{1*}, Maryam Gholamniya Foumani² and Hossein Falahzadeh¹

¹ Department of Biostatistics, Faculty of Health, Yazd Diabetes Research Center, Shahid Sadoughi University of Medical Sciences, Yazd, Iran

² School of Nursing and Midwifery, Guilan University of Medical Sciences, Rasht, Iran

* **Correspondence:** Email: Irsps.biostatistics@gmail.com.

Abstract: *Background:* The aim of this study was to compare the effectiveness of Cox model and Exponential parametric, Weibull, Log Normal and Log Logistic models in evaluating factors affecting retinopathy diagnostic time in patients with type 2 diabetes. *Methods:* In this prospective historical study, 400 patients with type 2 diabetes without retinopathy referred to the Ophthalmology Clinic of Yazd Diabetes Research Center in 2008 were followed up for diagnosis of retinopathy by January 2013. Significant variables in the univariate model were introduced into the Cox multivariate and parametric models to determine the effective factors on the time of retinopathy diagnosis. The criterion for comparing the performance of the models was the Akaike's criterion. All calculations were performed using R software and a significant level of 0.05 was considered. *Results:* The mean and median time of retinopathy diagnosis was 52.46 and 58 months, respectively. 3% of patients in less than one year and 16% of patients in less than two years of retinopathy were diagnosed. *Conclusion:* According to Akaike's criterion, Cox model has the best fit in determining the factors affecting the time of retinopathy diagnosis.

Keywords: diabetes; diabetic retinopathy; Cox model; parametric models; Akaike's criteria

1. Introduction

Survival analysis is a statistical process that has been profoundly influenced by scientific research and applied statistics in recent years. This field involves studying the survival time from a

defined and precise point to an incident or outcome of our interest; that is, so called failure [1]. The studied subject may be a person in a clinical trial or an animal in a veterinary study or a plant in a laboratory test. In many studies, the desired event or failure, is the occurrence of the event and in most cases is the death of the individual, but, it is not always a bitter accident [2]. Failure may be the improvement of the patient after a period of treatment, the relapse of an illness, the time elapsed until the first delivery after the marriage and etc. The main difference between survival analysis model with other models, such as linear regression or logistic regression, is the type of desired response variable as well as the existence of censorship. In survival analysis, the response variable is the time to event, and during the follow-up period censorship may occur, while in linear regression, the response is generally a continuous variable (such as blood pressure) [2], also, in the logistic regression, the response is a dual variable (such as a disease, its existence or absence), and in the case of censorship, it is necessary to remove the individual and can not use all the available information, but in the survival analysis, all the information of an investigated individual will be used until the last moment in the study [1]. In most cases, distribution of survival time is in our interest, but in general, our focus and attention in survival studies is the modeling of survival time as a response variable on a number of explanatory variables and their effects on survival time. Survival analysis and its models have grown widespread in recent years. The history of the development of survival analysis is based on the study of Cox in 1972 on seminal data and the introduction of a proportional hazard model [3]. Subsequently, parametric, nonparametric and semi-parametric methods for survival studies were introduced depending on the type of data and the type of censorship encountered in the study. The Cox model (proportional hazard model) emphasizes that the explanatory variables have a constant multiplier effect on the risk and base hazard function [4]. When variables are independent of time, this condition will result in a constant relative risk over time. In other words, the risk of people in different groups of treatment or people with different risk factors will have a constant ratio. The Cox model is the most practical method to assess the association of explanatory variables with survival response variable, but has its own limitation such as the estimation of proportional hazards. Nowadays, due to an increasing use of survival analysis in medical studies, a need for an efficient and flexible model for analysis of survival is severely felt. Most medical studies use Cox regression model for assessing the survival distribution of cancer patients, while alternative parametric models including Exponential models, Weibull, Log Normal, and, Log Logistic may be useful in some conditions [4]. The aim of this study was to compare the effectiveness of Cox model and Exponential, Weibull, Log Normal, Log Logistic parametric models in evaluating factors affecting retinopathy diagnostic time in patients with type 2 diabetes.

The prevalence of diabetes in Iranian adults aged 25–64 is 7.7%, which is higher than global statistics. Diabetes mellitus type-2 is characterized through pathophysiologic abnormalities including muscle, hepatic and adipocyte insulin resistance which lead to reduced glucose uptake, excessive glucose production and increased lipolysis. Furthermore, it causes beta cell failure and apoptosis, glucagon sensitivity of hepatocytes, reduced incretin effect, upregulated glucose production and increased reabsorption of glucose by renal tubular [5]. Unlike developed countries that have high-incidence of diabetes, older people in the developing world (between the ages of 46 and 64) develop diabetes, which adds to the burden of diabetes on developing societies [3,5]. It has been shown that living in rural and greenspace areas have been linked to reduced risk for diabetes, while, living in crowded cities because of noise pollution exposure followed by disrupted sleep patterns have been proven to be associated with increased risk for the disease [6]. Diabetes and its mortality

are closely related to socioeconomic status (SES) due to its prevalence in deprived groups [7]. Previous studies demonstrated that mortality risk of type 2 diabetes is doubled in patients from developing countries compared to developed countries population [8].

Type 2 diabetes patients are more susceptible to cardiovascular diseases such as ischemic stroke [9]. Diabetic retinopathy (DR) is a chronic, potentially sight-threatening complication of diabetes mellitus, which is a very specific microvascular disease of type 1 and type 2 diabetes, and increases the risk for ischemic stroke [10]. DR is one of the major causes of loss of vision [5] which is caused by retinal neovascularization, intraocular hemorrhages and finally tractional retinal and macular edema detachment. This event may lead to irreversible blindness. Individuals suffering from diabetes for more than 20 years usually experience DR resulted vision loss [11]. So, it seems necessary to investigate the rate of survival and association of independent variables with the time of retinopathy diagnosis using an efficient model such as Cox proportional hazard model, as well as Exponential, Weibull, Log Normal, Log Logistic parametric models and comparing these models to find the most efficient one in determining the affecting factors on retinopathy diagnosis time in type 2 diabetes mellitus patients.

2. Materials and methods

This study is a historical cohort study that evaluates the time of diagnosis of retinopathy in type 2 diabetic patients referred to Yazd diabetes research center from 2008 to 2013 and the effects of risk factors on this survival time.

2.1 Sampling method

Four hundred individuals were selected to be entered in this study based on the following formula at type one error of $\alpha = 0.05$ and the exponent of $1 - \gamma = 0.9$, $n = ((z_{1 - \alpha/2} + z_{1 - \gamma})/\beta)^2 \times (1/(p(1-p)(1-R^2)(1-\pi)))$. β is the estimate of the random variable coefficient U_i . U_i is a random variable with binomial distribution. Considering $\pi = 4.0$, $R^2 = 85.0$, $\beta = 1/1$, $p = 6.0$, $Z_{1 - \alpha/2} = 96.1$, $z_{1 - \gamma} = 28.1$, 400 patients with type 2 diabetes without retinopathy were selected.

2.2 Inclusion and exclusion criteria

Inclusion criteria for pre-existing retinopathy and exposure to any diabetes therapy (except for insulin) were defined for the exclusion criteria for the departure of patients from the study.

2.3 Data collection

Patients medical record were investigated and individual and clinical data including age, gender, duration of diabetes, smoking, familial history of diabetes, history of hypertension, fasting blood glucose, triglyceride, cholesterol, body mass index, insulin intake, diet habits, using oral pills, using aspirin pill, anemia, laser eye record in the first-degree relatives and using any medications beside insuline were retrieved. Besides, dilated fundus examination was performed on the patients in order to eliminate the possibility of pre-existing retinopathy. Then the patients were followed up for retinopathy diagnosis by January 2013. The average follow-up period from the first referral to the

end of the study was 46 months. The desired outcome, or the same failure, in this study was the diagnosis of retinopathy in the patients. The failure time is also the time to refer to the diagnosis of retinopathy by month. In the current study, censored cases include those who have not been diagnosed with retinopathy at the end of the study, as well as those who are missing in follow-up.

2.4. Data analysis

All the analysis were performed using statistical software R. After collecting the data, coding them, information is entered into the software R, and the Cox model and parametric models of Weibull, Exponential, Log Normal, Log Logistic were used to investigate the factors affecting the time of retinopathy diagnosis, which is the main objective of this study. The efficiency of the models has been compared with the Akaike's criterion. Also, the parametric models have been investigated using Cox-Vanell fit method.

3. Results

Of the 400 patients, 138 were male (34.5%) and 261 were female (65.5%). The mean age of the subjects was 54.73 ± 6.815 years. The average and mean retinopathy time for these patients was 46.517 and 58 months, respectively. 3% of patients in less than 1 year and 16% of patients in less than 2 years, and 22% of patients in less than 3 years and 29% of patients in less than four years of retinopathy were diagnosed. Of the 400 patients under study, 152 (38%) patients suffered from retinopathy and 248 (62%) patients were considered censored observations (retinopathy for them not detected or missing in the study) (Table 1). None of the investigated patients had used any diabetes therapy except for insulin.

Based on the results of single-variable analysis shown in Table 2, age, BMI, cholesterol, sex, fasting blood glucose and triglyceride are not effective on the time of retinopathy diagnosis ($P < 0.05$). Except that the variable duration of diabetes was significant ($P < 0.05$).

Table 1. The mean of the variables studied.

Variables	Mean \pm SD
Age	54.73 \pm 6.81
Duration of diabetes (years)	5.74 \pm 7.06
Fasting blood sugar	9.187 \pm 3.63
Triglyceride	5.224 \pm 8.128
Cholesterol	198 \pm 3.46
BMI	54.22 \pm 58.3
To diagnose retinopathy	72.11 \pm 3.6
Duration between referrals	52.46 \pm 08.17
To diagnose retinopathy	

The relationship between independent variables and retinopathy diagnostic time by applying parametric, Exponential, Weibull, Log Normal, and Log Logistic is demonstrated in Table3.

Table 2. Single-variable analysis of factors affecting the time of retinopathy.

Characteristic	Cox		
	HR	(95%CI)	p
Age (year)	1.023	(0.998,1.049)	0.065
BMI	0.976	(0.93,1.025)	0.34
Cholesterol	0.999	(0.995,1.004)	0.797
Duration of diabetes	1.088	(1.069,1.07)	0.000
Gender (male = 1, female = 0)	1.027	(0.736,1.432)	0.877
Fasting blood sugar	1.003	(0.999,1.006)	0.087

Table3. The relationship between independent variables and retinopathy diagnostic time by applying parametric, Exponential, Weibull, Log Normal, Log logistic models.

Characteristic	Exponential			Weibull			Log normal			Log logistic		
	HR	95%CI	P*	HR	95%CI	P*	HR	95%CI	P*	HR	95%CI	P*
Age (year)	1.022	0.997, 1.046	0.078	1.017	0.998, 1.036	0.064	1.018	0.999, 1.037	0.056	1.0195	1.1, 0.0391	0.051
BMI	0.977	0.929, 1.025	0.36	0.982	0.946, 1.018	0.33	0.98	0.942, 1.017	0.30	0.979	0.941, 1.017	0.29
Cholesterol	0.999	0.995, 1.003	0.81	0.999	0.996, 1.002	0.8	0.999	0.996, 1.002	0.75	0.999	0.996, 1.002	0.79
Duration of diabetes	1.074	1.058, 1.089	0.00	1.058	1.046, 1.071	0.00	1.086	1.064, 1.108	0.00	1.095	1.069, 1.121	0.00
Gender (male = 1, female = 0)	1.03	0.697, 1.363	0.86	1.02	0.767, 1.274	0.87	0.979	0.705, 1.253	0.88	0.992	1.266, 0.718	0.96
Triglyceride	0.999	0.998, 1.001	0.64	0.999	0.998, 1.00	0.62	0.999	0.998, 1.0007	0.46	0.999	0.998, 1.0008	0.53
Fasting blood sugar	1.002	0.999, 1.005	0.10	1.001	0.99, 1.004	0.08	1.001	0.99, 1.004	0.13	0.002	0.99, 1.004	0.099

P*: chi2 test, significant at the level of 0.05; HR: hazard rate

Based on the Table 4, the best model among the parametric and Cox models is the Cox proportional hazard model with the least acoustic value.

Table 4. The most efficient way of results based on the Akaike's values.

Model	Akaike's value
Cox	1504.553
Exponential	1548.798
Weibull	1526.146
Log Normal	1512.651
Log Logistic	1515.214

3. Discussion

So far, many studies have been carried out on the use of the Cox regression model, but based on a systematic study, only 5% of them have considered the appropriateness of the risks [12]. As described in detail earlier, we have implemented our statistical models based on retinopathy

diagnostic data of 400 patients with type 2 diabetes. It was found that the basis of selection among the models was the selection of proportional risk model. Cox model, had the least amount of Akaike among all models (AIC = 1504/553). The assumption that the Cox risks were fitted using a good fit method showed that this hypothesis exists for all variables in the final model of Cox. Among parametric models, the Log Normal model and Log Logistic had better fit to these data. From the clinical point of view, and the topic of diabetic retinopathy and the study of the factors influencing it, extensive studies have been carried out in this field in our country and other countries [13]. The risk of retinopathy in men is 2% higher in women than in women. In other words, the time to diagnose retinopathy in men is 2% earlier than women. But, this difference was not statistically significant, and in similar studies, gender has not been identified as an effective predictor of retinopathy [14]. In a study by Koohtan et al. (2012) that examined the prevalence and factors associated with diabetic retinopathy, despite higher prevalence of diabetes in women (74.6%) than men (25.4%), no significant statistical relationship between sex and retinopathy was found [14]. In this study, BMI was not recognized as a risk factor for retinopathy. In the Manaviat study, which was performed on 120 type 2 diabetic patients in Yazd, there was no significant relationship between BMI and retinopathy [15]. In the study conducted by Zhang et al. the prevalence of diabetic retinopathy in the United States had a significant correlation between the body mass index and diabetic retinopathy [16]. Diagnosis and treatment of anemia in diabetic retinopathy management is very important. P.K. Rani (2010) found that anemia increased the risk of diagnosis of diabetic retinopathy by 80% compared to those who did not have anemia [17]. In Francisco study, anemia was a factor in the development of diabetic retinopathy [18]. In a prospective cross-sectional study in India containing 306 patients with type 2 diabetes was found that anemia affects diabetic retinopathy [19]. In the present study, the risk of diagnosis of diabetic retinopathy in people with anemia was 1.4% fold greater than those without anemia. In this study, there was no significant relationship between triglyceride, cholesterol, fasting blood glucose, smoking, and retinopathy, which is similar to another study in this area [15]. In the research of Agha Dostouli and Sadr (2004) who examined the prevalence and risk factors of diabetic retinopathy in diabetic patients, the results showed that cholesterol and triglyceride levels did not had a statistically significant correlation with the incidence of diabetic retinopathy [20]. In a study by Shafi Pour et al, on 540 diabetic patients in Sari (2006), did not show a relationship between diabetic retinopathy and smoking [21] which is consistent with the results of this study. In our study, the duration of diabetes was recognized as a risk factor for retinopathy and also as a risk factor for the diagnosis of retinopathy, so that one year increase in duration the risk of developing retinopathy would be 9% in the Cox model. In other words, with the increase in the duration of diabetes, the time to diagnose retinopathy is reduced. In the Bin He study, the duration of diabetes is correlated with diabetic retinopathy [22]. The Bueso et al study found that the risk of retinopathy in patients with 5–10 years of diabetes was doubled, and in people who had over 15 years of diabetes were 4.5 times greater than those with less than 5 years of age [23]. Also, in our study, there was no history of diabetes in the family as a risk factor for retinopathy, and similar studies did not show a significant relationship between the incidence of retinopathy and the history of diabetes in the family [13]. According to the results of this study and in comparison with other studies, factors including anemia, aspirin use, insulin use and duration of diabetes were associated with retinopathy. Regarding the relationship between retinopathy and the factors mentioned, follow up and examine these patients by internal medicine specialist and ophthalmologists is often recommended to control these factors and to raise awareness of patients and to diagnose and control

retinopathy. An initial examination for early diagnosis of diabetic patients with retinopathy and identifying and controlling the risk factors of this disorder is important in increasing the survival rate of patients, reducing the economic burden of retinopathy and improving the quality of life. Statistically, there are also studies on the use of parametric and semi-parametric models in the analysis of survival data. Rajaifard et al. (2009) examined the application of parametric models of survival analysis in gastric cancer. The results of the data analysis were approximately the same with the Cox model and the parametric models. Based on the Akaike's scale, the Weibull and Exponential models had the best fit for survival data with the values of 848 and 850, respectively [24]. The study by Grover et al. (2012) entitled "Parametric Method for Estimating the Time of Survival of Diabetic Nephropathy with Censored Data, Survival Time (Diabetic Diagnosis Time to Nephropathy)" had a Weibull Parameter Distribution. Also, fitting the model using Cox-Snell was investigated, the lines corresponding to the Cox-Snell residuals were close to the straight line and the Weibull model was introduced for fitting the data [25]. In our study, the corresponding lines of Cox-Snell residues were close to the straight line in the Log Normal and, Log Logistic models, and these parametric models were introduced to fit the data of retinopathy diagnostic time. Bradburn et al. (2003) examined the adequacy of a variety of parametric models and Cox proportional hazards through the remainder and the Akaike's criteria. Orbe (2002) studied the fitting of parametric and semi-parametric models in gastric cancer. The results showed that the Log Normal model had the best fit. Also, the findings showed that the assumption of the proportionality of risks for using the Cox model was not present in this study [26]. Giolo et al. (2012) in a study on patients with heart failure found that the data were slightly distorted and that the Cox model was not suitable for fitting data [27]. That does not match the results of our study. In our study, the assumption of the proportionality of risks for the Cox model was established and the Cox model was the most efficient model.

4. Conclusion

In the current study, the assumption of the proportionality of risks for the Cox model was established and the Cox model introduced as the most efficient model. Because, the AIC standard was lower than other models, and this model was better fitted to the data, and this model gives us more accurate results. Although the Cox model estimates are more familiar to medical field researchers, the results from parameter model estimates that are expressed as relative risk are not very well known for these researchers and their interpretation is similar to the Cox model's risk level (HR). These models are easily estimated using the maximum likelihood method, and they are able to examine linear, nonlinear effects and provide researchers with more acceptable results when the risk assumption is not appropriate. However, for more accurate comparisons of these models, the effects of sample size and percentage censorship are needed. In this study, it was determined that the assumption of the proportion of hazard risks is accepted for the semi-parametric Cox model. Although, the Cox model risk ratio and parametric models were approximately the same, based on the Akaike's criterion, the Cox model has a more accurate fit to the time of retinopathy diagnosis, and also based on the Akaike's criterion and the Cox-Snell model charts of the Log Normal model and, Log Logistic has the best fit between parametric models and can be used as a substitute for the Cox model in analyzing the factors affecting retinopathy diagnostic time in patients with type 2 diabetes.

Acknowledgment

This manuscript was extracted from an MSc thesis approved by council of Shahid Sadoughi University of Medical Sciences, Yazd, Iran. We appreciated the staff of council of Shahid Sadoughi University of Medical Sciences.

Conflict of interest

The authors declare no conflict of interest.

References

1. Kleinbaum D (2004) *Survival Analysis a Self-Learning Text*. New York: Springer.
2. Salehi M, Vazirinasab H, Khoshgam M, et al. (2012) Application of the generalized additive model in determination of the retinopathy risk factors relation types for Tehran diabetic patients. *Razi J Med Sci* 97: 1–9.
3. Gaafar M, Khattab A (2013) Prevalence and Predictors of Diabetic Retinopathy Among Elderly type II diabetics. *Am Sci* 9: 639–646.
4. Kostev K, Rathmann W (2013) Diabetic retinopathy at diagnosis of type 2 diabetes in the UK: a database analysis. *Diabetologia* 56: 109–111.
5. DeFronzo RA (2018) Pathogenesis of type 2 diabetes mellitus. *Diabetes Epidemiol Genet Pathog Diagn Prev Treat*: 181–253.
6. Mackenbach JD, den Braver NR, Beulens JWJ (2018) Spouses, social networks and other upstream determinants of type 2 diabetes mellitus. *Diabetologia* 61: 1517–1521.
7. Walker J, Colhoun H, Livingstone S, et al. (2018) Type 2 diabetes, socioeconomic status and life expectancy in Scotland (2012–2014): a population-based observational study. *Diabetologia* 61: 108–116.
8. Walker JJ, Livingstone SJ, Colhoun HM, et al. (2011) Effect of Socioeconomic Status on Mortality Among People With Type 2 Diabetes. *Diabetes Care* 34: 1127.
9. Hjelmgren O, Strömberg U, Gellerman K, et al. (2019) Does retinopathy predict stroke recurrence in type 2 diabetes patients: A retrospective study? *PloS One* 14: e0210832.
10. Groeneveld Y, Tavenier D, Blom JW, et al. (2019) Incidence of sight-threatening diabetic retinopathy in people with Type 2 diabetes mellitus and numbers needed to screen: a systematic review. *Diabetic Med*.
11. Srinivasan R (2018) Diabetes mellitus: Can retinopathy be far behind? *J Curr Res Sci Med* 4: 78.
12. Ngandu NH (1997) An empirical comparison of statistical tests for assessing the proportional Hazards Assumption of Cox's Model. *Stat Med* 16: 611–626.
13. Askarishahi M, Hajizadeh E, Afkhami-Ardekani M (2011) Factors affecting retinopathy in patients with type 2 diabetes by analyzing the current status data. *Tehran Univ Med J* 68: 674–680.
14. Kohian H, Sohrabi MB, Zolfaghari P, et al. (2013) The Prevalence and Related Factors of Diabetic Retinopathy in Shahroud. *Knowl Health J* 7: 179–184.

15. Manaviat M, AfkhamiArdekani M, Rashidi M (2006) Assessment four-year incidence of retinopathy and its risk factors in type II diabetes. *J Shaeed Sdoughi Univ Med Sci Yazd* 14: 11–19.
16. Zhang X, Saaddine J, Chou C, et al. (2010) Prevalence of Diabetic Retinopathy in the United States, 2005–2008. *J Am Med Assoc* 304: 649–656.
17. Rani PK, Raman R, Rachepalli SR, et al. (2010) Anemia and Diabetic Retinopathy in Type 2 Diabetes Mellitus. *J Assoc Physicians India* 58: 91–94.
18. Francisco J, Patricia P, Martha G, et al. (2012) Anemia as a factor related to the progression of proliferative diabetic retinopathy after photocoagulation. *J Diabetes Complications* 26: 454–457.
19. Diep TM, Tsui I (2013) Risk factors associated with diabetic macular edema. *Diabetes Res Clin Pract* 100: 298–305.
20. Aghadoust D, Sadr F (2005) Prevalence and risk factors of diabetic retinopathy in diabetic patients referred to Kashan Diabetes center in 1381-1382. *Q J Fyz* 33.
21. Shafiepour L, Ahmadzade A, Shafiepour V, et al. (2006) Prevalence of diabetic retinopathy and associated risk factors among Patients referred to Hospital Imam Khomeini Surrey city. *J Mazandaran Univ Med Sci* 16: 125–130.
22. He BB, Wei L, Gu YJ, et al. (2012) Factors Associated with Diabetic Retinopathy in Chinese Patients with Type2 Diabetes Mellitus. *Int J Endocrinol*.
23. Bueso SE, Vigo SJ, Pereze F, et al. (2005) Prevalence of diabetic retinopathy in the regional community of Extremadura 1997-2001. *Arch Soc Esp Oftalmol* 80: 187–189.
24. Rajaeefard A, Dehkordi B, Tabatabaee H, et al. (2009) Applying parametric models for survival analysis of gastric cancer. *Arch SID* 13: 83–88.
25. Rover G, Sabharwal A (2012) A Parametric Approach to Estimate Survival time of diabetic nephropaty with left truncated and right censored data. *Int J stat Probab* 1: 128–137.
26. Orbe J, Ferreira E, Nunez-Anton V (2002) Comparing proportional hazards and accelerated failure time models for survival analysis. *Stat Med* 21: 493–510.
27. Giolo SR, Krieger JE, Mansur AJ, et al. (2012) Survival Analysis of Patients with Heart Failure Implications of Time-Varying Regression Effects in Modeling Mortality. *PloS One* 7: 1–8.



AIMS Press

© 2019 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)