What Is the Best Parametric Survival Models for Analyzing Hemodialysis Data?

Mohsen Vahedi1, Mahmood Mahmoodi1, Kazem Mohammad1, Sharzad Ossareh2 & Hojjat Zeraati1

1 Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, IR Iran
2 Nephrology Section, Hasheminejad Kidney Center, Iran University of Medical Sciences, Tehran, IR Iran

Correspondence: Hojjat Zeraati, Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Pursina St., Keshavarz Blvd., Tehran 14155-6446, IR Iran. Tel: 98-21-8898-9126. Fax: 98-21-8898-9127. E-mail: zeraatih@tums.ac.ir

Received: January 4, 2016   Accepted: January 27, 2016   Online Published: February 24, 2016
doi:10.5539/gjhs.v8n10p118          URL: http://dx.doi.org/10.5539/gjhs.v8n10p118

Abstract

Background: Chronic kidney disease (CKD) and end-stage renal disease (ESRD) are both common public health problems worldwide. Hemodialysis (HD) is one of the main ultimate modalities of renal replacement therapy in these patients. The aim of this study was to compare the different parametric (Weibull, Gamma, Gompertz, Log-logistic and Lognormal) survival models, in maintenance HD (MHD) patients.

Method: This study was conducted from March 2004 to October 2013 and encompassed 544 ESRD patients under MHD in Hasheminejad Kidney Center, Tehran, Iran. Laboratory, clinical and demographic data were extracted from the Hemodialysis Data Processor Software, which had been designed for data collection in Hasheminejad Kidney Center. Exponential, Weibull, Gompertz, lognormal and log-logistic were used for analyzing survival of hemodialysis patient using STATA software. To compare these models Akaike Information criterion (AIC) and Cox-Snell residual were utilized.

Results: According to the both criteria (AIC and Cox-Snell residual), Weibull survival model manifested better results as compared with other models. According to this model, age at the time of admission (HR=1.015, p-value=0.018), walking ability (HR=0.656, p-value=0.010), diabetes mellitus as the underlying disease (HR=1.392, p-value=0.038), hemoglobin level (HR=0.790, p-value<0.001), serum creatinine (HR=0.803, p-value<0.001), serum protein (HR=0.747, p-value=0.010) and Single pool Kt/V (HR=0.092, p-value<0.001), had significant effect on survival of the hemodialysis patient.

Conclusion: In our analysis Weibull distribution, which had the lowest AIC value, was selected as the most suitable model.

Keywords: survival analysis, hemodialysis, weibull distribution, log-logistic distribution, lognormal distribution, gamma distribution, gompertz distribution

1. Introduction

Chronic kidney disease (CKD) and end-stage renal disease (ESRD) are two common worldwide public health problems in recent years (Modi & Jha, 2006; Yang et al., 2015). Hemodialysis (HD) is one of the main modalities of renal replacement therapy in ESRD patients (Mousavi, Hayati, Valavi, Rekabi, & Mousavi, 2015), however, despite great improvements in HD machines, techniques and dialyzers, the mortality of maintenance HD (MHD) is still high, compared to that of the general population (Port, 1994). The annual mortality rate of MHD patients was reported from 9% and 16% in Japan and Europe, respectively to 24% in the United States(Foley, Parfrey, & Sarnak, 1998). Age of commencement dialysis and diabetes mellitus as the underlying cause of ESRD are shown as the main determinants of high mortality in MHD patients, together with the presence of certain comorbidities such as cardiovascular disease, hypertension and low serum albumin at the start of dialysis (Ahmed, Dimitrov, Perna, Remuzzi, & Nahas, 2009; Shibiru, Gudina, Habte, Derbew, & Agonafer, 2013; Sikole et al., 2007).

The epidemiology of ESRD in Iran showed that the prevalence and incidence of ESRD have been increasing in Iran, from 238 case per million population (pmp) and 49.9 pmp, respectively in 2000 to 357 pmp and 63.8 pmp,
respectively in 2006 (Aghighi et al., 2008). Hence, like other parts of the world ESRD is becoming a major public health problem and needs a good health care strategy. Investigating the survival of these patients is an important health issue.

Survival analysis models for investigating the effect of covariates on the risk (hazard) function of a population is a favorable statistical technique in the last decade. Generally there are two different kinds of models for survival analysis: non parametric models including Cox proportional hazard (PH) model and parametric survival accelerated failure time (AFT) models including Weibull, Exponential, Log-logistic, Lognormal and Gamma models (Kleinbaum & Klein, 2012).

Although cox model is a popular one for analyzing survival data, in some conditions due to the characteristics of the data, parametric models are better than cox model. The most important characteristic is data distribution, which could be linked to a special parametric model. Also with a decrease sample size, relative efficiencies may further changed in favor of parametric models. When empirical information is sufficient, parametric models can provide some insight into the shape of the baseline hazard (Efron, 1977; Nardi & Schenper, 2003; Ng'andu, 1997; Oakes, 1983). Therefore, using the best model to fit the data is important in survival analysis.

It may be mentioned that there are very few studies in Iran regarding dialysis practice and survival data. Hence, the aim of this study was to use and compare the different parametric (Weibull, Gamma, Gompertz, Log-logistic and Lognormal) models for the analysis of the hemodialysis patients.

2. Method

This was a retrospective cohort study from March 2004 to October 2013 in HD ward of Hasheminejad Kidney Center, Tehran, Iran. The clinical, demographic and laboratory data of 544 incident and prevalent HD patients were extracted from Hemodialysis Data Processor Software, which was designed for hemodialysis data collection, at Hasheminejad Kidney Center, a main referral kidney hospital in Iran. The Demographic data were collected at admission. Laboratory tests were performed in each month, or every two or six months. The variables including age at admission, Gender(Males/Females), Marital status(Married, living with partner/Single, divorced, widowed Married), Smoking(Yes/No), Walking ability(Walks without help/Walks with help, Uses wheelchair or crutches, Unable to walk), diabetes mellitus(Yes/No), hypertension(Yes/No), hemoglobin level, plasma levels of calcium, phosphate, intact parathyroid hormone (PTH), potassium, triglyceride, low-density lipoprotein cholesterol (LDL), high-density lipoprotein cholesterol (HDL), creatinine and protein and single-pooled Kt/V (SpKt/V), as a measurement of efficacy of dialysis. Finally the mean of each variable was calculated for the analysis.

Survival analysis included two basic models of parametric and nonparametric. Two models were used for correction of the effect of covariates on the survival function: accelerated failure time (AFT) and proportional hazard (PH). Parametric survival models are statistically more powerful than non-parametric or semi-parametric models. They are AFT models and this models use natural logarithm of the survival period ln (t) is explained by a linear function on of the covariates. Exponential, Weibull, Gompertz, lognormal and log-logistic were used as the different parametric models. We used these different parametric models for analysis survival of hemodialysis patients, first in univariate and then in multivariate analysis. For comparison among models, two different criteria were used; Akaike Information criterion (AIC) and Cox-Snell residual (Kleinbaum & Klein, 2012).

For AIC, the lower value and for Cox-Snell residual, the line more close to bisector of the first quarter, were used as the indicator of the best model. All statistical analyses were done using the Stata Statistical Software: Release 11. College Station, TX: Stata Corp LP. P-values less than 0.05 were considered significant.

3. Results

A total of 544 patients participated this study, 236 (42.4%) of whom were female. The mean (±SD) age at admission time was 56.03 (±17.03) years, and ranged from 11 to 89 years. The median survival of the patients was 56 months. The median of follow up time for dead patient was 27.76 month and for censors was 24.73 month. 1, 3, 5 and 9 year survival rates were 0.92, 0.66, 0.46 and 0.25 respectively.

At the end of the study 216 patients (38.8%), had died. The cause of death in this study for patient are as follow: Cardiovascular 98 (45.6%), Cerebrovascular 27 (12.6%), Gastrointestinal 4 (1.9%), Infectious 30 (13.9%), Neoplastic 13 (6.0%), Pulmonary 6 (2.8%), Trauma or accident 1 (0.5%), Not identified 36 (16.7%).

The patients, who died or were transplanted, transferred to another center or recovered, were considered as right censored. Table 1 showed the characteristics of all patients.

Figure 1 shows the Cox-Snell residuals for these 5 models. If the hazard function follows the 45 degree line then
we know that the model fits the data well.

Table 1. Characteristic of patients on hemodialysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>Censor</th>
<th>Dead</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (Year)</td>
<td>51.5±17.38</td>
<td>62.71±13.9</td>
<td>55.93±16.99</td>
</tr>
<tr>
<td>Gender (Males)</td>
<td>188(57.15)</td>
<td>128(59.54)</td>
<td>316(58.09)</td>
</tr>
<tr>
<td>Marital status (Married, living with partner)</td>
<td>256(77.82)</td>
<td>175(81.40)</td>
<td>431(79.23)</td>
</tr>
<tr>
<td>Smoking (Yes)</td>
<td>59(17.93)</td>
<td>42(19.53)</td>
<td>101(18.57)</td>
</tr>
<tr>
<td>Walking ability (Walks without help)</td>
<td>282(85.71)</td>
<td>135(62.79)</td>
<td>417(76.65)</td>
</tr>
<tr>
<td>Diabetes mellitus (Yes)</td>
<td>95(28.87)</td>
<td>99(46.04)</td>
<td>194(35.66)</td>
</tr>
<tr>
<td>Hypertension (Yes)</td>
<td>161(48.93)</td>
<td>92(42.79)</td>
<td>253(46.51)</td>
</tr>
</tbody>
</table>

| Variable                          | Exponential HR P>|z| | Gompertz AIC | Weibull HR P>|z| | Log-logistic AIC | LogNormal AIC |
|-----------------------------------|------------------|------------------|---------------|------------------|------------------|
| Age (Years)                       | 1.031 0.000 929.268 1.033 0.000 | 923.896 1.033 0.000 | 906.548 1.025 0.000 | 900.038 1.042 0.000 | 896.312 |
| Male                              | 1.248 0.121 972.238 1.268 0.088 | 969.699 1.293 0.065 | 953.058 1.216 0.089 | 939.461 1.230 0.077 | 935.278 |
| Married, living with partner      | 1.237 0.224 973.123 1.251 0.202 | 970.934 1.265 0.180 | 954.624 1.231 0.146 | 940.212 1.226 0.164 | 936.447 |
| Smoking                           | 1.205 0.279 973.537 1.252 0.195 | 971.031 1.282 0.150 | 954.536 1.133 0.378 | 941.588 1.108 0.484 | 937.912 |
| Walks without help                | 0.410 0.000 938.569 0.385 0.000 | 932.581 0.376 0.000 | 914.195 0.472 0.000 | 904.496 0.468 0.000 | 901.006 |
| Diabetes mellitus                 | 1.827 0.000 955.861 1.986 0.000 | 949.517 2.023 0.000 | 931.593 1.597 0.000 | 925.341 1.538 0.000 | 924.853 |
| Hypertension                      | 0.908 0.483 974.175 0.924 0.566 | 972.306 0.932 0.612 | 956.250 0.903 0.376 | 941.569 0.876 0.257 | 937.113 |
| Dialysis frequency                | 0.881 0.600 974.400 0.901 0.664 | 972.455 0.925 0.746 | 965.406 0.974 0.901 | 942.340 0.994 0.977 | 938.399 |
| Hemoglobin (g/dl)                 | 0.852 0.003 965.938 0.835 0.001 | 961.911 0.818 0.000 | 943.486 0.862 0.001 | 930.658 0.880 0.001 | 927.587 |
| Serum creatinine (mg/dl)          | 0.770 0.000 906.734 0.758 0.000 | 899.876 0.747 0.000 | 878.394 0.801 0.000 | 861.120 0.805 0.000 | 857.263 |

Note: Data were presented as means ± standard deviation or frequency (percentage); PTH: Parathyroid hormone; LDL: Low-density lipoprotein; HDL: High-density lipoprotein; SpKt/V: Single pool Kt/V.

Based on both criteria (AIC and Cox-Snell residual), Weibull survival model had the best fit compared to other parametric models, following with lognormal, log-logistic and Gompertz, and the worst model was exponential (Table 2 and Table 3). Based on the Weibull model, age on the admission time (HR=1.015, p-value=0.018), walking ability (HR=0.656, p-value=0.010), Diabetes mellitus (HR=1.392, p-value=0.038), hemoglobin (HR=0.790, p-value=0.001).

Table 2. Results of parametric models for univariate analysis in hemodialysis patients
Table 3. Results of parametric models for multivariate analysis in hemodialysis patients

<table>
<thead>
<tr>
<th></th>
<th>Exponential</th>
<th>Gompertz</th>
<th>Weibull</th>
<th>Log-logistic</th>
<th>LogNormal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (Year)</td>
<td>1.014</td>
<td>1.017</td>
<td>1.015</td>
<td>1.006</td>
<td>1.007</td>
</tr>
<tr>
<td>Male</td>
<td>1.125</td>
<td>1.116</td>
<td>1.119</td>
<td>1.230</td>
<td>1.202</td>
</tr>
<tr>
<td>Married, living with partner</td>
<td>0.801</td>
<td>0.802</td>
<td>0.837</td>
<td>0.891</td>
<td>0.882</td>
</tr>
<tr>
<td>Smoking</td>
<td>1.077</td>
<td>1.204</td>
<td>1.230</td>
<td>1.054</td>
<td>1.070</td>
</tr>
<tr>
<td>Walks without help</td>
<td>0.717</td>
<td>0.659</td>
<td>0.656</td>
<td>0.714</td>
<td>0.714</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>1.295</td>
<td>1.441</td>
<td>1.392</td>
<td>1.174</td>
<td>1.142</td>
</tr>
<tr>
<td>Hypertension</td>
<td>0.856</td>
<td>0.911</td>
<td>0.913</td>
<td>0.861</td>
<td>0.826</td>
</tr>
<tr>
<td>Dialysis frequency</td>
<td>1.042</td>
<td>1.142</td>
<td>1.149</td>
<td>1.091</td>
<td>1.179</td>
</tr>
<tr>
<td>Hemoglobin (g/dl)</td>
<td>0.869</td>
<td>0.818</td>
<td>0.790</td>
<td>0.878</td>
<td>0.890</td>
</tr>
<tr>
<td>Serum creatinine (mg/dl)</td>
<td>0.830</td>
<td>0.815</td>
<td>0.803</td>
<td>0.853</td>
<td>0.846</td>
</tr>
<tr>
<td>Serum potassium (meq/L)</td>
<td>0.927</td>
<td>0.923</td>
<td>0.883</td>
<td>0.912</td>
<td>0.888</td>
</tr>
<tr>
<td>Serum calcium (mg/dl)</td>
<td>0.927</td>
<td>0.895</td>
<td>0.891</td>
<td>0.936</td>
<td>0.934</td>
</tr>
<tr>
<td>Serum phosphate (mg/dl)</td>
<td>0.987</td>
<td>0.994</td>
<td>0.986</td>
<td>0.989</td>
<td>0.990</td>
</tr>
<tr>
<td>Serum intact PTH (mg/dl)</td>
<td>1.000</td>
<td>1.000</td>
<td>0.999</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>LDL (g/dl)</td>
<td>1.002</td>
<td>1.003</td>
<td>1.004</td>
<td>1.002</td>
<td>1.002</td>
</tr>
<tr>
<td>HDL (g/dl)</td>
<td>0.983</td>
<td>0.983</td>
<td>0.983</td>
<td>0.991</td>
<td>0.989</td>
</tr>
<tr>
<td>Serum triglyceride (mg/dl)</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Serum protein (g/dl)</td>
<td>0.809</td>
<td>0.761</td>
<td>0.747</td>
<td>0.889</td>
<td>0.887</td>
</tr>
<tr>
<td>SpKt/V</td>
<td>0.163</td>
<td>0.110</td>
<td>0.092</td>
<td>0.245</td>
<td>0.254</td>
</tr>
</tbody>
</table>

Note: HR: Hazard Ratio; RR: Relative Risk; P>|z|: p-value; AIC: Akaike Information Criterion; PTH: Parathyroid hormone; LDL: Low-density lipoprotein; HDL: High-density lipoprotein; SpKt/V: Single pool Kt/V.
Figure 1. Cox-snell residual plot for multivariate parametric models

Serum creatinine (HR=0.803, p-value<0.001), serum protein (HR=0.747, p-value=0.010) and SpKt/V (HR=0.092, p-value<0.001), had significant effect on survival of the hemodialysis patient.

4. Discussion

The objective of this study was to know the best survival model for analyzing of the hemodialysis patients. It was a single center study with 544 prevalent and incident maintenance hemodialysis patients. We presented the results of 9 years patient survival and risk factors of mortality during a mean of follow-up of 64.19 months.

In our study, age at admission was one of the most important contributors to patient mortality in MHD patients. With each one year increased in age, patient mortality increased by 1.5%. This is similar to other studies which indicated that mortality of older patients at admission is greater than younger ones (Sa Carvalho, Henderson, Shimakura, & Sousa, 2003). Age is also a major risk factor for adverse outcomes of peritoneal dialysis (Ahmad & Shahzad, 2015; Coric et al., 2015; Mailloux et al., 1994; Tsai et al., 2013). But in the other study on elder people, there is no difference between survival of elderly and “older “patient (Jeloka, Sanwaria, Periera, & Pawar, 2016). From all these studies, we can conclude young age can affected survival of this patient.

Diabetes mellitus was another risk factor in our models, and had a significant effect on patient survival. According to our results, having diabetes mellitus increases the hazard of mortality by 39% in MHD patient. Some previous studies confirm this finding (Schiller et al., 2015; Sikole et al., 2007; Vavallo et al., 2014).

Walking ability was another factor that interestingly had a positive effect on survival of MHD patients in our study and we found significant lower hazards of death in those who could walk without help. A recent study has shown the correlation between gait speed and mortality in HD patients (Inaba et al., 2013). They rationalized that as walking challenges the heart, lungs, circulatory, nervous, and musculoskeletal systems, gait speed provides an informative marker of the overall health status (Kutner, Zhang, Huang, & Painter, 2015).

Our model indicated the protective effect of serum creatinine on patient survival consistent with previous studies (Dwyer et al., 2005; Moreau-Gaudry et al., 2011). Serum creatinine has a considerable correlation with lean body mass (Dwyer et al., 2005). So this protective effect may be due to better maintenance of body mass and good nutrition (de Souza, Matos, Barros, & Rocha, 2014; Moreau-Gaudry et al., 2011).

Mineral metabolism affects mortality in hemodialysis patients and is identified by imbalances in calcium, serum
phosphate and parathyroid hormone (PTH) (Taniguchi et al., 2013; Tentori et al., 2008). Based on our findings, serum phosphate, calcium and PTH levels had no significant effect on mortality of these patients. All these finding have been confirmed by other studies (Inaba et al., 2013; Lin et al., 2015).

Moreover, Hemoglobin and HDL level were other predictors of mortality in our study. The positive effect of high level of both factor are confirmed by this study and other studies (Coric et al., 2015; Moradi et al., 2014; Teixeira, Lopes, Silva, & Santos, 2015; Tsubakihara, Akizawa, Iwasaki, & Shimazaki, 2015).

Efficacy of dialysis has been shown as the predictor of survival in various studies. The clear relationship between low dialysis efficacy and patient's survival and the high mortality has been observed previously (Lertdumrongluk et al., 2014; Ramirez et al., 2012; Singh et al., 2013). In addition, in our results, clear positive effect of high level of SpKt/V on survival of ERSD patient was confirmed.

In this study, we used parametric models for analyzing the survival of ERSD patients under MHD. In such survival analyses, usually researchers use proportional Cox models (Royston, 2004; Therneau & Grambsch, 2000). This model needs some assumptions such as proportionality of hazards among various variables. If these assumptions are violated, Cox models are not appropriate and using parametric models will be the best option. These parametric models assume a specific distribution for time variable in survival analysis and fit the model to the data, without a need to the Cox model pre-assumptions (Collett, 2003; Hougaard, 2000). The parametric models include lognormal, log-logistic, Weibull, Gamma and Gompertz distributions. Lognormal distribution was introduced by McAlister in 1897 (Aktürk Hayat et al., 2010). The skewed distribution, where the average value is low and variance is high, is the main characteristics of this distribution (Aktürk Hayat et al., 2010). Cancer survival studies like chronic leukemia, which is positively skewed, are analyzed via lognormal distribution (Lee & Wang, 2003). Log-logistic distribution is for the random variables, which is positive in probability and statistics. Many researchers in different fields such as breast cancer and HIV studies, used this model in survival analysis (Byers et al., 1988; Gupta, Akman, & Lvin, 1999; Zhou, Mi, & Guo, 2007). Where the mortality ratio increases at the very beginning and decrease at the end this model is recommended (Lee & Wang, 2003).

Weibull distribution is a generalised version of the exponential distribution. It is a flexible distribution that allows a monotonous increase and an important field in which this model is used (Viscomi et al., 2006). Gamma distribution has an adaptive characteristic that makes it appropriate for survival models (Aktürk Hayat et al., 2010). Survival in chronic hepatitis, and in patients with nasopharyngeal tumor are the examples of studies that have used this parametric model (Bolin & Greene, 1986; Galli, Maini, Salvatori, & Andreasi, 1983; Poon et al., 2004).

Gompertz model is another parametric model that has been frequently used in modelling the mortality ratio data by medical researchers. Relation with tumor development in different cancer studies is an example of using this model (Ahuja & Nash, 1967).

Limitations of our study were single center and the low time of follow up (just 9 years). On the other hand we couldn’t add alcohol consumption one of the risk factor of ERCP patient, in our analysis. Information about this factor is not possible due to some cultural problems.

5. Conclusion

In this study, we aimed to compare the results of the survival analysis of ERSD patient using Weibull, Gamma, Gompertz, Log-logistic and Lognormal models. In conclusion, by using AIC, the models obtained via Weibull, Log-logistic, Lognormal, Gamma and Gompertz models were compared and the most suitable model for MHD survival analysis was specified. Although the AIC values of the five distributions were very close to each other, the Weibull distribution, which had the lowest AIC value, was selected as the most appropriate model. So it is conclude that the Weibull distribution is the best model for survival analysis of MHD patients.

Acknowledgments

This paper is a part of a research for a PhD degree at School of Public Health and Institute of Public Health Researches, Tehran University of Medical Sciences (TUMS), Iran, and was financially supported by the university. The authors thank the staff of the hemodialysis unit of Hasheminejad Kidney Center for their great distribution in providing the data.

Conflict of Interest

The authors declare that there is no conflict of interests regarding the publication of this paper.
References


Copyrights
Copyright for this article is retained by the author(s), with first publication rights granted to the journal.
This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).