

# Parametric Survival Models of Hemodialysis Patients in Relation with Patient-Related Factors

Reem Yousif MEKKI<sup>1,2</sup>, Mohamed Hassan MUDAWI<sup>4</sup>, Manahil Saidahmed MUSTAFA<sup>1</sup>,  
Altaiyb Omer Ahmed MOHMED<sup>1</sup>, Ahmed Bakheet Abd ALLA<sup>3</sup>, Abdel Rahman AHMED<sup>5</sup>

## Abstract

Background: Survival analysis refers to analyzing of statistical data for which the outcome variable of interest is time until an event occurs. This research aimed at comparing different models of parametric Proportional Hazards (PH) models (Weibull, exponential, Gompertz) in patients with hemodialysis to determine the best model for assessing the survival of patient. Study consists of 325 hemodialysis patients who referred to public hospitals in Khartoum state in the period from December 2005 to December 2015. Data was used to estimate the survival function with view to identify risk factors influencing among end-stage renal disease (ESRD) population. Based on the Cox-Snell Residuals and AIC, BIC, and Gompertz (PH) model is an efficient model than other when the values of (AIC=662.21), (BIC=703.83) and (R<sup>2</sup>=0.211) where maintained Study assessed that the variables dealing with univariate models were significant but had a significant effect on hemodialysis survival. The Gompertz model had the smallest AIC and BIC value; therefore; it was selected as the most appropriate model. In multivariable analysis, the BIC had the lowest value and the highest value in each analysis. The study assessed that diabetes mellitus and hypertension, regular, and hospital, had a significant effect.

**Keywords:** survival analysis, parametric models, hemodialysis, Cox-Snell residuals, AIC, BIC, R<sup>2</sup>.

## INTRODUCTION

Survival analysis focuses on estimating the probability about individual who will hazard for a given length of time until death. Survival analysis is particularly useful when the probability of occurrence of the event under study changes with time<sup>1-3</sup>.

The final stage of chronic kidney disease is End-stage Renal Disease (ESRD) and is characterized by pro-

gressive permanent kidney failure. Dialysis therapy is a procedure aimed at eliminating the body's excrement and toxic substances and compensating for the loss of function of the kidneys. One dialysis class is hemodialysis<sup>4</sup>. It has been estimated that more than 1.1 million patients worldwide are estimated to have ESRD, with an addition of 7 percent annually. For example, incidence and prevalence levels in the United States are projected to increase by 44% and 85%, respectively,

<sup>1</sup> College of Science, Sudan University of Science and Technology

<sup>2</sup> College of Computer Science and Mathematics, University of Bahri, Sudan

<sup>3</sup> College of Medical Laboratory Science, Sudan University of Science and Technology

<sup>4</sup> College of Social and Economic Studies, University of Bahri, Sudan

<sup>5</sup> College of Petroleum Geology and Minerals, University of Bahri, Sudan

### Corresponding author:

Reem Yousif MEKK, College of Science, Sudan University of Science and Technology.

E-mail: reem11229@hotmail.com

from 2000 to 2015, and incidence and prevalence rates per million population by 32% and 70%. The progress of ESRD patients in developing countries has similar trends<sup>5</sup>.

Sudan is one of the countries where the chronic kidney failure is alarming. The frequency reported rate of (ESRD) new cases in Sudan is 70-140 per million people annually. The data available about the root cause of renal disease that leads to chronic renal disease is very limited<sup>5,6</sup>.

Scientific studies have uncovered major causes of end stage renal disease in survival time. These causes are affecting in the survival of hemodialysis patients for a live long time. Millions of people are being affected with outbreak of kidney disease around the world.

This study is to compare the performance of different parametric models of the survival of hemodialysis patients. Parametric models were selected to estimate the survival probabilities. The application of these models helps to identify the prognostic factors that resulted in increasing the probability of survival.

## MATERIALS AND METHODS

This study consisted of 325 hemodialysis patients who referred to public hospitals named Ahmed Gasim, Ibn Sina, Omdurman, Selma center, Bahri, and Ribat in Khartoum State during the period of time (December 2005 to December 2010) and then they were followed-up till 2015. Data captured age, date of diagnosis of the disease, survival status even in the case of death or alive per months, sex, marital status, education level and occupation.

### DATA COLLECTION

Khartoum State composed of 3 biggest cities named Khartoum, Bahri and Omdurman. Data of the study were collected from the biggest and well-known public hospitals in these three cities. Total number of patients covered in our study was 325.

### INCLUSION CRITERIA

In the period from December 2005 to December 2015, all hemodialysis patients referred to the 6 public hospitals were included and all age ranges were included.

### EXCLUSION CRITERIA

Patients with hemodialysis that have been diagnosed with acute renal failure, inadequate medical history, pa-

tients with hemodialysis who have stayed for a brief period of time and those in emergency conditions have been removed.

## DATA ANALYSIS

The descriptive statistical analysis, percentage, and frequency were measured using Microsoft Excel software. In addition to the variables considered in this analysis, qualitative variables like (sex, marital status, education, occupation, daily dialysis, weekly frequency dialysis, hospital, diabetes mellitus and hypertension, diabetes mellitus, polycystic kidney disease, renal obstructions, shrunken kidney, unknown and other) and quantitative variables were classified into effective variables.

The log-rank test is a statistical test used to compare the survival distributions of two or more groups used to test the hypothesis where there is no difference between the categories for and variable. It does not provide any estimation of the actual size of the effect; in other words, it provides a statistical, but not a clinical, assessment of the effect of the factor.

In this analysis, the quantitative variables were not distributed as usually calculated by the Kolmogorov-Smirnov test when the probability value was higher than the significance level of 0.05, so the parametric method was used.

The form of Survival Analysis has been applied in this study was heavily relied on both a Univariate where there is only one explanatory variable required and Multivariate where at least two explanatory variables 7 of patients with chronic kidney disease diagnosed with ESRD under hemodialysis care as they explained in Tables 5 and 6.

### Parametric models

A parametric survival model time is supposed to follow a certain distribution, which its probability density function can be represented by unknown parameters. Weibull, Exponential, Gompertz, log logistic, lognormal and gamma distributions are widely used<sup>1,8,9</sup>.

### Parametric proportional hazard (PH) Models

Cox (1972) introduced the parametric (PH) model it's also known as the Cox regression model. The widely used models are Exponential, Weibull and Gompertz distribution<sup>2,10,11</sup>.

### Exponential Distribution

The simplest and most important distribution in survival studies is the exponential distribution. It is often referred to as a purely random failure pattern.

The hazard function is

$$h(t) = \lambda, \quad t \geq 0 \quad (1)$$

A constant, independent of  $t$ .

The corresponding survivorship function is

$$S(t) = e^{-\lambda t} \quad t \geq 0 \quad (2)$$

and so the implied The probability density function Survivorship function is

$$f(t) = \lambda e^{-\lambda t} \quad t \geq 0, \quad \lambda > 0 \quad (3)$$

A high  $\lambda$  value shows high risk and limited survival; a low  $\lambda$  value shows low risk and long survival. The distribution is also referred to as the unit exponential when  $\lambda = 1^{2,3,11}$ .

Under the exponential PH model, the hazard function of a particular patient with covariates  $x_p, x_2, x_3, \dots, x_p$  is given by

$$h(t|x) = \lambda \exp(B_1x_1 + B_2x_2 + B_3x_3 + \dots + B_px_p) = \lambda \exp(\hat{B}x)$$

### Weibull Distribution

Proposed by Weibull (1939) and its applicability to different cases of failure, again discussed by Weibull (1951). In several studies of reliability and mortality from human diseases, it was then used<sup>2,3,11</sup>.

$$h(t) = \lambda \gamma t^{\gamma-1}, \quad (4)$$

Amore general form of hazard function is such that

The survivor function is

$$S(t) = \left\{ - \int_0^t \lambda \gamma u^{\gamma-1} du \right\} = \exp(-\lambda t^\gamma) \quad (5)$$

The corresponding probability density function is then

$$f(t) = \lambda \gamma t^{\gamma-1} \exp(-\lambda t^\gamma), \quad (6)$$

The shape and scale parameters are therefore called gamma and  $\lambda$ , under the Weibull PH model, the hazard function of a specific patient is provided by the hazard function of a specific patient.

$$h(t|x) = \lambda \gamma t^{\gamma-1} \exp(B_1x_1 + B_2x_2 + B_3x_3 + \dots + B_px_p) = \lambda \gamma t^{\gamma-1} \exp(\hat{B}x)$$

### Gompertz Distribution

The Gompertz model has found application in demography and the biological sciences. In the particular case where  $\gamma=0$ , the hazard function has a constant value. The hazard function increases with time, decreases with time<sup>2,11</sup>.

$$h(t) = \exp(\lambda + \gamma t) \quad (7)$$

This shows that linear int. is the log-hazard function. Monotonically, the Gompertz risk increases or decreases. The survival function is

$$S(t) = \exp \left[ -\frac{e^\lambda}{\gamma} (e^{\gamma t} - 1) \right], \quad (8)$$

And the corresponding density function is

$$f(t) = \exp \left[ (\lambda + \gamma t) - \frac{1}{\gamma} (e^{\lambda + \gamma t} - e^\lambda) \right], \quad (9)$$

Under the Gompertz PH model, the hazard function of a particular patient is given by

$$h(t|x) = \lambda \exp(\gamma t) \exp((B_1x_1 + B_2x_2 + B_3x_3 + \dots + B_px_p)) = \lambda \exp(\hat{B}x) \exp(\gamma t)$$

### Selection Criterion

One of these criteria is the information criterion of Akaike (AIC), the Bayesian Information Criterion (BIC) and the Cox-Snell Information Criterion (CSIC), the latter of which is a graphic rather than a mathematical criterion, many of the criteria used to choose the best model from different models deal with the same data for prediction in the future.

**AIC:** Comparisons may also be made on the basis of statistics between a variety of potential models which do not necessarily need to be nested<sup>2,12-14</sup>.

$$AIC = -2(\log \text{likelihood}) + 2(P + K) \tag{10}$$

Where P is the number of parameters, and K is the number of (excluding constant) coefficients in the model. For P=1, for P=2, for Weibull and Gompertz, for the exponential. The smaller the value of this statistic, the better the model, the better this statistic is known as Akaike’s knowledge criterion.

**BIC:** The Bayesian Information Criteria (BIC) is given by<sup>15,16</sup>.

$$BIC = -2(\log \text{likelihood}) + (P + K) * \log(n) \tag{11}$$

In the distribution, where P is the number of parameters, K is the number of coefficients and n is the number of observations. As the best-fit model, the distribution that has the lowest BIC value is considered a metric to test the goodness-of-fit of a regression model for proportional hazards. As a descriptive statistics for goodness of fit, Hosmer and Lemeshow propose the following<sup>17,18</sup>.

$$R_p^2 = 1 - \left\{ \exp \left[ \frac{2}{n} (L_0 - L_p) \right] \right\} \tag{12}$$

Where the log likelihood for the fitted model with p covariates is  $L_p$ , and  $L_0$  is the log likelihood for model zero, the model without covariates is the log probability.

**Cox Snell-Residual:** The Cox-Snell Residuals can be used to check the quality of the model’s fit<sup>15,19</sup>, defined as the person with observed time  $t$

$$r_{ci} = \hat{H}_i(t_i) = -\log \hat{S}_i(t_i), \tag{13}$$

Where  $H_i(t_i)$  is estimated cumulative hazard function, and  $S_i(t_i)$  is the estimated survivor function in equation are plotted against<sup>1</sup>.

### Ethical Considerations

The study protocol was authorised by the ethics and research committees of the Ministry of Health of Khartoum (serial number: KMOH-REC-1-2020). Hospitals received informed consent.

## RESULTS

Our findings showed that a total of 325 patients with hemodialysis were enrolled in this study. The demographic characteristics of the targeted patients showed that 59.7 % were male, 40.3 % were female in terms of sex. By December 2015, 52.3 % of patients had died and 47.7 % were still alive, according to survival status. The marital status of the patients showed that 2.5% were divorced, 71.4% were married, 24% were single and 2.2% were widowed. Education revealed that 7.7 % of patients were illiterate, 32.6 % received basic education, 4 % were intermediate, 39.1 % completed secondary education and 16 % graduated. Patients’ occupation wise shows that 18.8 % were employees, 13.8 % were freelancers, 41.2 % were unemployed, 3.7 % were police officers, 4.3 % were retired 7.4 % were students, 11.8 % were professionals.

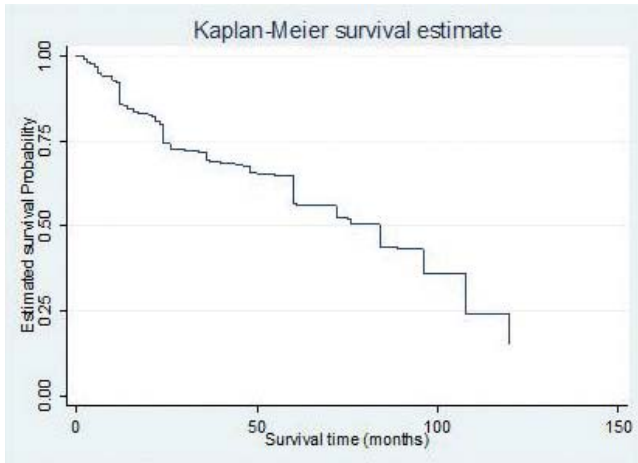
In regard to the qualitative variables such as age; the minimum age was 6years. The maximum age was 88years. The first quartile was 46.03years. The median age was 45years. While third quartile was 75years. The results breakdowns were as presented in Table 1.

Results of clinical characteristics showed that 88.9 % of patients with hemodialysis were normal and 11.1 % were sporadic patients with hemodialysis, 27.4 % were diabetic mellitus and 72.6 % were not diabetic mellitus. 29.5 % had hypertension and 70.5 % had no hypertension. 89.8% had both diabetes mellitus and hypertension, and 10.2% had neither diabetes mellitus nor hypertension. 3.4 % had shrunken kidneys and 96.6 % had no shrunken kidneys. Dialysis frequency per week found that two times (8.8%) and three times (81.2%) had polycystic kidney disease and 94.8% had no polycystic kidney disease. 8.0 % had renal obstruction and 92.0 % had no renal obstruction. 9.5 % were uncertain and 90.5 % were uncertain. 5.8 % had each other, and 94.2 % had no other.

Based on the log rank test, the variables considered to be important in Table 1 and Table 2 with p-value > 0.05 were entered in the mean parametric model, while other variables were not significantly excluded from the parametric model. The variables used in the parametric model were normal, dialysis frequency per week, hospitals, diabetes mellitus, hypertension, diabetes mellitus and hypertension, shrunken kidneys, other.

The median overall survival time was estimated at 84 months and the trust level was found at 95% (61-89) as shown in Figure 1, which clarified the overall survival curve of hemodialysis patients.

For univariate analysis, an additive Weibull and Gompertz model are found with similar meaningful



**Figure 1.** Kaplan-Meier to estimate survival function for hemodialysis patients.

findings for all Wald test variables ( $P$ -value  $< 0.05$ ). Based on hazard ratio factors; age, diabetes mellitus, diabetes mellitus + hypertension, increased risk of death and other variables were observed; regular, hospital, hypertension, kidney shrunk dialysis frequency per week, and other, respectively, variables were found to have substantially higher survival rates. (see Table 3, 4, and 5).

Based on a multivariate analysis, it was assessed that risk factors, including age, hospital, dialysis frequen-

cy per week, daily dialysis, were significant relative to other variables for the Wald test ( $P$ -value  $< 0.05$ ). The most important factors in hemodialysis patients were HR variables, including age (1.014), diabetes mellitus (1.127), diabetes mellitus + hypertension (1.165). On the other side, other factors, such as regular (0.581) hospital, have been noted (see Table 3, 4, and 5).

Table (6) displays AIC, BIC and  $R^2$  for different models: Exponential, Weibull and Gompertz. Gompertz has the lowest AIC (662.21), BIC (703.83), and the highest  $R^2$  (0.211).

The variables that match the univariate parametric model are shown in Table 7. We also found that the Gompertz model is the safest one to use in the future for forecasting. It is selected as it has the lowest AIC and BIC value.

For these three models, Figures 2, 3 and 4 displays the Cox-Snell residuals, the cumulative hazard function of Cox-Snell residuals (vertical axis) against the Cox-Snell residuals (horizontal axis) calculated below in Map. The fitness of the survival model is more fitting for the short deviation of residuals from the straight line through the origin with a slope of 1. Then, based on criteria (AIC, BIC) and residual Cox-Snell, the Gompertz model is the better model compared to another parametric model.

**Table 1.** Demographic characteristics

Variable	Categories	Frequency N= 325	Percentage %	Log-rank test	
				Chi-square test	P-value
<b>Sex</b>	male	194	59.70%	0.66	0.416
	female	131	40.30%		
<b>Marital status</b>	divorced	8	2.50%	4.92	0.1778
	married	232	71.40%		
	single	78	24%		
	widowed	7	2.20%		
<b>Education status</b>	illiterate	25	7.70%	0.5	0.9732
	basic	106	32.60%		
	intermediate	15	4.60%		
	secondary	127	39.10%		
	graduates	52	16%		
<b>Occupation</b>	employee	61	18.80%	7.3	0.294
	freelancers	45	13.80%		
	unemployed	134	41.20%		
	policemen	11	3.70%		
	retired	14	4.30%		
	student	24	7.40%		
	worker	36	11.1.8%		
<b>Age</b>	minimum	maximum	first quartile	median	third quartile
	6	88	46.03	45	75

Table 2. Clinical characteristics

Variable	Categories	Frequency N= 325	Percentage%	Log-rank test	
				Chi-square test	P-value
daily dialysis	irregular	36	11.10%	5.35	0.000
	regular	289	88.90%		
Hospital	AhemdGasim	50	15.40%	88.57	0.000
	Bhari	70	21.50%		
	EbnSena	74	22.80%		
	Omdurman	62	19.10%		
	Ribat	25	7.70%		
	Salma Center	44	13.50%		
<b>Causes of end stage renal failure among study population</b>					
DIABETES MELLITUS	No	236	72.60%	7.21	0.007
	Yes	89	27.40%		
HYPERTENSION	No	229	70.50%	20.31	0.000
	Yes	96	29.50%		
DIABETES MELLITUS AND HYPERTENSION	No	292	89.80%	4.7	0.0302
	Yes	33	10.20%		
SHRUNKEN KIDNEYS	No	314	96.60%	6.39	0.0115
	Yes	11	3.40%		
DIALYSIS FREQUENCY PER WEEK	Two times	61	18.80%	21.65	0.000
	three times	264	81.20%		
POLYCYSTIC KIDNEY DISEASE	No	308	94.80%	0.05	0.8305
	Yes	17	5.20%		
RENAL OBSTRUCTIONS	No	299	92.00%	0.13	0.7167
	Yes	26	8.00%		
UNCERTAIN	No	294	90.50%	1.25	0.2634
	Yes	31	9.50%		
OTHER	No	306	94.20%	7.63	0.0058
	Yes	19	5.80%		

other=Systemic lupus erythematosus, tropical disease (malaria), Gout, cardiovascular disease, NSAID.

Table 3. Analysis results for hemodialysis patients using an exponential parametric model in univariate and multivariate analysis

Variable	Univariate			Multivariate		
	Coef ( $\hat{\beta}$ )	HR[exp( $\hat{\beta}$ )]	p-value	Coef ( $\hat{\beta}$ )	HR[exp( $\hat{\beta}$ )]	p-value
Age	0.021	1.021	0.000	0.014	1.014	0.012
Daily dialysis	-0.987	0.373	0.000	-0.543	0.581	0.017
Hospital	-0.225	0.798	0.000	-0.162	0.850	0.004
Diabetes mellitus	0.41	1.507	0.010	0.120	1.127	0.564
Hypertension	-0.423	0.655	0.024	-0.414	0.661	0.076
Diabetes mellitus +hypertension	0.833	2.300	0.000	0.153	1.165	0.578
Shrunken kidneys	-2.061	0.127	0.04	-1.804	0.165	0.075
Dialysis frequency per week	-0.408	0.665	0.024	-0.409	0.664	0.018
Other	-1.076	0.341	0.018	-0.849	0.428	0.075
Intercept				-2.419		0.001

Coef; coefficient, HR;Hazard Ratio, p=value significant at < 0.05 level of significance

Table 4. Analysis results for hemodialysis patients using a Weibull parametric model in univariate and multivariate analysis

Variable	Univariate			Multivariate		
	Coef ( $\hat{\beta}$ )	HR[exp( $\hat{\beta}$ )]	p-value	Coef ( $\hat{\beta}$ )	HR[exp( $\hat{\beta}$ )]	p-value
Age	0.021	1.021	0.000	0.014	1.014	0.012
Daily dialysis	-1.046	0.351	0.000	-0.562	0.570	0.014
Hospital	-0.242	0.785	0.000	-0.179	0.836	0.001
Diabetes mellitus	0.415	1.514	0.009	0.121	1.129	0.562
Hypertension	-0.418	0.658	0.026	-0.426	0.653	0.069
Diabetes mellitus +hypertension	0.878	2.406	0.000	0.133	1.142	0.632
Shrunken kidneys	-2.077	0.125	0.038	-1.830	0.160	0.071
Dialysis frequency per week	-0.736	0.479	0.000	-0.388	0.678	0.026
Other	-1.115	0.328	0.014	-0.876	0.417	0.066
Intercept				-3.438		0.000
Shape parameter				1.249	50	

Coef; coefficient, HR; Hazard Ratio, p=value significant at < 0.05 level of significance

Table 5. Analysis results for hemodialysis patients using a Gompertz parametric model in univariate and multivariate analysis

Variable	Univariate			Multivariate		
	Coef ( $\hat{\beta}$ )	HR[exp( $\hat{\beta}$ )]	p-value	Coef ( $\hat{\beta}$ )	HR[exp( $\hat{\beta}$ )]	p-value
Age	0.022	1.022	0.000	0.015	1.015	0.009
Daily dialysis	-1.032	0.356	0.000	-0.514	0.598	0.025
Hospital	-0.251	0.776	0.000	-0.191	0.826	0.001
Diabetes mellitus	0.415	1.514	0.009	0.128	1.136	0.542
Hypertension	-0.406	0.666	0.031	-0.411	0.663	0.079
Diabetes mellitus + hypertension	0.900	2.460	0.000	0.137	1.147	0.623
Shrunken kidneys	-2.072	0.126	0.039	-1.831	0.160	0.071
Dialysis frequency per week	-0.726	0.484	0.000	-0.359	0.698	0.040
Other	-1.137	0.321	0.012	-0.876	0.416	0.066
Intercept				-3.021		0.000
Shape parameter				.011		0.000

Coef; coefficient, HR; Hazard Ratio, p=value significant at < 0.05 level of significance, other= (Systemic lupus erythematosus, tropical disease (malaria), Gout, cardiovascular disease, NSAID).

Table 6. Scores of Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) and R2 for multivariate parametric models

Models	Exponential	Weibull	Gompertz
AIC	678.95	669.47	662.21
BIC	716.78	711.09	703.83
R <sup>2</sup>	0.198	0.209	0.211

## DISCUSSION

This research compared various parametric (PH) models to determine the best model for assessing and analyzing the risk factors affecting patients with hemodialysis survival in public hospitals in Khartoum State. In this analysis, we closely tracked the medical history of the targeted patients in the hospitals through the

duration before the occurrence of a significant event such as death or living.

In the analysis of survival results, the focus is always on the probability or risk of death at any time after the initial period. One of the reasons for modeling data on survival is to decide which combinations of possible explanatory variables especially affect the type of the hazard function, the care that causes the risk of death

Table 7. Scores of Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) for Univariate Parametric Models

Variable	Exponential		Weibull		Gompertz	
	AIC	BIC	AIC	BIC	AIC	BIC
Age	715.34	722.91	709.45	720.8	702.65	714
Daily dialysis	717.67	725.27	710.66	722.01	705.01	716.37
Hospital	711.96	719.53	703.17	714.52	694.46	705.81
Diabetes Mellitus	728.48	736.01	723.06	734.41	717.01	728.36
Diabetes Mellitus +hypertension	729.29	736.86	724.13	735.48	718.4	729.75
Hypertension	721.78	729.35	715.26	726.61	708.58	719.93
Shrunken kidneys	725.46	733.03	719.96	731.31	713.98	725.33
Dialysis Frequency per week	716.65	724.21	711.29	722.64	705.69	717.04
Other	726.74	734.31	720.76	732.11	714.3	725.66

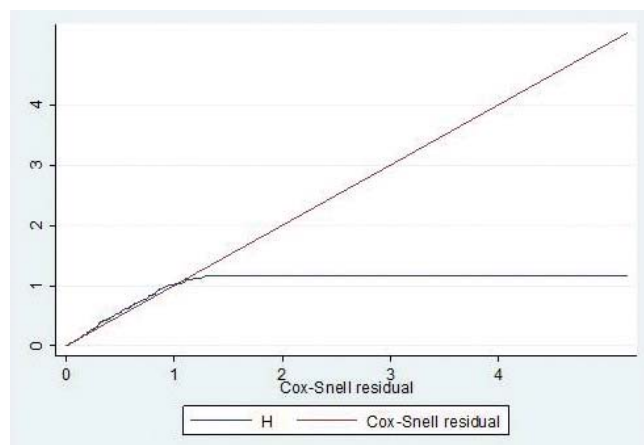


Figure 2. Cox-Snell plot for multivariate Exponential parametric model.

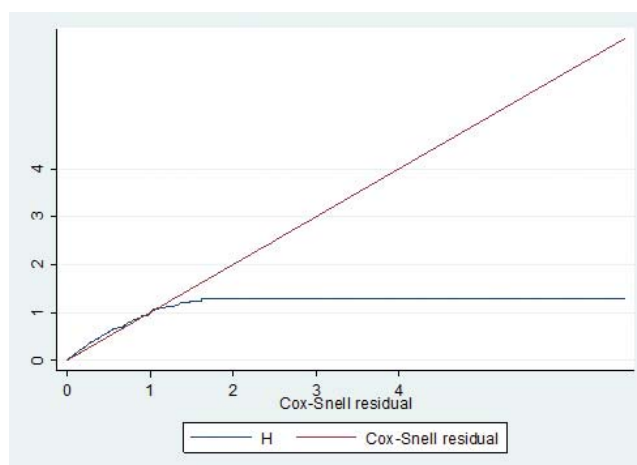


Figure 4. Cox-Snell plot for multivariate Gompertz parametric model.

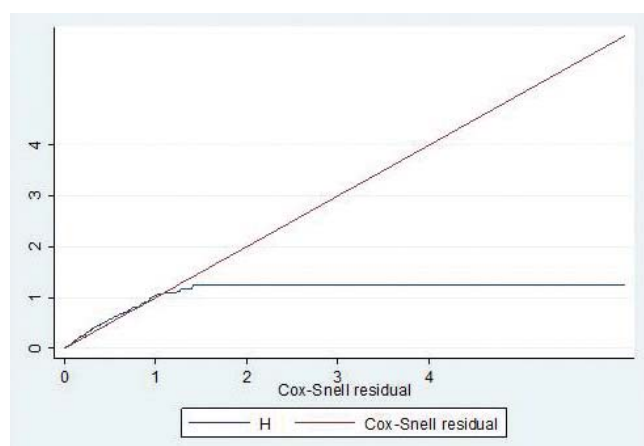


Figure 3. Cox-Snell plot for multivariate Weibull parametric model.

can be measured and the degree to which the hazard feature is impaired by other explanatory variables can be assessed. Another justification for modeling the ha-

zard function is for individuals to achieve an approximation of the hazard function itself<sup>2</sup>.

The quantitative variables were not distributed as usual in this research, as indicated in the data analysis, so the parametric models were used. In the log-rank test, the variables were important and were incorporated into parametric models. To estimate variables, univariate and multivariate tests were used.

The univariate analysis study for three models (Exponential, Weibull, and Gompertz) showed that all variables were important effects. We also found that age, diabetes mellitus, both diabetes mellitus and hypertension, increased the risk of death in patients (shorter survival) so that they could influence survival in the univariate model of this study. Other variables (regular, hospital, hypertension, shrunk kidneys, dialysis frequency per week, others) have decreased the risk of death (longer survival) and have a direct effect on the



survival of the hemodialysis patient.

Multivariate analysis showed that many variables were insignificant. From the results based on Criteria (AIC, BIC) and the highest  $R^2$ . Multivariate analysis found that many variables were negligible. From the results based on Parameters (AIC, BIC) and the highest  $R^2$  in addition to Cox-Snell residual, we have found that Gompertz is the best model. It is therefore the most efficient fit model among other parametric models for patient hemodialysis data in addition to Cox-Snell residual; we found that the Gompertz is the best model.

Similar findings were found in previous studies<sup>20</sup>, indicating that age, and diabetes mellitus were important variables and hypertension, and the frequency of dialysis per week was negligible. Reference to (Exponential, Weibull and Gompertz) univariate and multivariate statistical analysis in hemodialysis patients showed that Weibull was chosen as the most effective model. The univariate findings were important in this analysis, but they were negligible in the multivariate. In another study conducted by 21, era, diabetes mellitus, hypertension, were negligible except for clinical significance. It was shown that the Weibull model was the best fit among the parametric models of hemodialysis patients. In this analysis, only the clinic and hospital variable was important in the Exponential model.

In view of the Gompertz multivariate model, which was a major value of the Wald test for various variables, like daily dialysis ( $\beta$ -pop=-0.514, HR =  $e^{\beta}$ -pop=0.598, P=0.025) hospital ( $\beta$ -pop=-0.191, HR =  $e^{\beta}$ -pop=0.826, P=.001), age ( $\beta$ -pop=0.015, HR =  $e^{\beta}$ -pop=1.015, P=0.009), dialysis Frequency per week ( $\beta$ -pop=-0.359, HR= $e^{\beta}$ -pop=0.698, P=0.040), when the (P<0.05) was achieved. That said, the approximate coefficient and risk ratio were also important and had a clear impact on the survival of hemodialysis patients in this study.

In the light of the above study, the model centered on these variables was important among the other models that were excluded from this research. These va-

riables (daily dialysis, hospital, age, dialysis Frequency per week) must therefore be included in the model. The final multivariate Gompertz PH model was then defined.

$$h(t|x) = 0.049 \exp(0.011 t) \exp(-0.514 \text{ daily dialysis} - 0.191 \text{ hospital} + 0.015 \text{ age} - 0.359 \text{ dialysis Frequency per week})$$

According to HR, the variables including age, diabetes mellitus, diabetes mellitus +hypertension, were considered to be highly significant factors in hemodialysis patients in the three models used in the research in particular multivariate analysis. Whereas other factors, such as regular in dialysis, hypertension, shrunken kidneys, dialysis frequency per week, and other, had substantially lower survival rates.

This research has its drawbacks, that is, the incompleteness of the majority of patient records and the lack of data that make it difficult to determine the real cause of the outbreak of the disease in Sudan. It is due to the fact that certain variables were not included in this analysis because they were not included in the patient medical record.

## CONCLUSION

Gompertz distribution model is being the best for hemodialysis patient's analysis. Some variables such as (age, daily dialysis, hospital, dialysis frequency per week) were significant factors. The study clarified that some variables like regular in dialysis were significant factor.

**Compliance with ethics requirements:** The authors declare no conflict of interest regarding this article. The authors declare that all the procedures and experiments of this study respect the ethical standards in the Helsinki Declaration of 1975, as revised in 2008(5), as well as the national law. Informed consent was obtained from all the patients included in the study.

## References

1. Kleinbaum, David G, and Mitchel Klein (2012). *Survival Analysis: A Self-Learning Text*. Third Edition. New York: Springer Science + Business Media.
2. Collett, David. (2003). *Modelling Survival Data in Medical Research*. Second Edition. London: Chapman & Hall.
3. Lee, Elisa T, and John Wenyu Wang.(2003). "Some Well-Known Parametric Survival Distributions and Their Applications", *Statistical Methods for Survival Data Analysis*. Third Edition. Canada. John Wiley & Sons, Inc., Hoboken, New Jersey.
4. Pagee, Jama Patient (2007). *Chronic Kidney Disease*. jama.jamanetwork.com (Accessed 6 28, 2012).
5. Elamin, Abdelsamee E. M., Nagah A. A. Mohammed, and Gad Allah Modawe (2012). Aetiology of End-Stage Renal Disease among Adult Sudanese Patients. *Sudan medical journal*,7(4):255-258.
6. Banaga, A. S., Mohammed, E. B., Siddig, R. M., Salama, D. E., Elbashir, S. B., Khojali, M. O., et al (2015). Causes of end stage renal failure among haemodialysis patients in Khartoum State/ Sudan. *BMC Res Notes*, 8:502.
7. Altman, DG, BL De Stavola, SB Love, and KA Stepniowska. (1995). Review of survival analyses published in cancer journals. *British Journal of Cancer*,72: 511-518.
8. George, Brandon, Samantha Seals, and Inmaculada Aban. (2014). Survival analysis and regression models. *J Nucl Cardiol*,21(4):686-694.
9. Abdelaal, Medhat Mohamed Ahmed, and Sally Hossam Eldin Ahmed Zakria. (2015). Modeling Survival Data by Using Cox Regression Model. *American Journal of Theoretical and Applied Statistics*,4(6):504-512.
10. Bradburn, M J, T G Clark, S B Love, and D G Altman.(2003). Survival analysis. Part II: Multivariate data analysis – an introduction to concepts and methods. *British Journal of Cancer*, 89(3):431-436.
11. Lawless, Jerald F. (2003). *Statistical Models and Methods for Lifetime Data*. Second Edition. New York: John Wiley & Sons.
12. Pourhoseingholi, Mohamad Amin, Ebrahim Hajizadeh, Bijan Moghimi Dehkordi, Azadeh Safaee, Alireza Abadi, and Mohammad Reza Zali.(2007).Comparing Cox Regression and Parametric Models for Survival of Patients with Gastric Carcinoma. *Asian Pacific J Cancer Prev*,8: 412-416.
13. Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans Automatic Control*,19(6):716-723.
14. Klein, John P, and Melvin L Moeschberger.(1997).*Survival Analysis: Techniques for Censored and Truncated Data*. New York: Springer-verlag.
15. Saikia, Rinku, and Manash Pratim Barman. (2017).A Review on Accelerated Failure Time Models. *International Journal of Statistics and Systems*,12: 311-322
16. Schwarz, Gideon E. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*,6: 461-46
17. Hosmer JR, David W, Stanley Lemeshow, and Susanne May. (2011). *Applied survival analysis: regression modeling of time to event data*. Hoboken, New Jersey: John Wiley & Sons,
18. Stevenson, Mark. December (2007). *An Introduction to Survival Analysis*. EpiCentre.massey.ac.nz.
19. Cox, D.R, and E.J Snell. E.(1968).A General Definition of Residuals (with discussion). *Journal of the Royal Statistical Society*, 30: 248-275.
20. Vahedi, Mohsen, Mahmood Mahmoodi, Kazem Mohammad, et al.:(2016). What is the Best Parametric Survival Models for Analyzing Hemodialysis Data. *Global Journal of Health Science*, 8(10):118-126.
21. Montaseri, Maryam, Jamshid Yazdani Charati, and Fateme Espahbodi (2016). Application of Parametric Models to a Survival Analysis of Hemodialysis Patients. *Nephrourol Mon*,8(6): e28738.