Survival Analysis of Patients with End Stage Renal Disease

To cite this article: J D Urrutia et al 2015 J. Phys.: Conf. Ser. 622 012014

View the article online for updates and enhancements.

Related content

- Present status and one upgrading method with a cold compressor of the EAST sub-cooling helium cryogenic system
  Zhigang ZHU, Qiyong ZHANG, Yuntao SONG et al.

- Observation on the transformation domains of super-elastic NiTi shape memory alloy and their evolutions during cyclic loading
  Xi Xie, Qianhua Kan, Guozheng Kang et al.

- Barrier layer non-uniformity effects in anodized aluminum oxide nanopores on ITO substrates
  Piao Liu, Vijay P Singh and Suresh Rajaputra
Survival Analysis of Patients with End Stage Renal Disease

J D Urrutia, W S Gayo, L A Bautista and E B Baccay
Faculty Members, Department of Mathematics and Statistics, College of Science, Polytechnic University of the Philippines, Sta. Mesa, Manila, Philippines
E-mail: jackieurrutia20@gmail.com, math_urrutia@yahoo.com.ph, jdurrutia@pup.edu.ph, willygayo@gmail.com, wsgayo@pup.edu.ph, labautista1971a@yahoo.com, labautista@pup.edu.ph, math_baccay@yahoo.com, ebbaccay@pup.edu.ph

Abstract. This paper provides a survival analysis of End Stage Renal Disease (ESRD) under Kaplan-Meier Estimates and Weibull Distribution. The data were obtained from the records of V. L. Makabali Memorial Hospital with respect to time t (patient’s age), covariates such as developed secondary disease (Pulmonary Congestion and Cardiovascular Disease), gender, and the event of interest: the death of ESRD patients. Survival and hazard rates were estimated using NCSS for Weibull Distribution and SPSS for Kaplan-Meier Estimates. These lead to the same conclusion that hazard rate increases and survival rate decreases of ESRD patient diagnosed with Pulmonary Congestion, Cardiovascular Disease and both diseases with respect to time. It also shows that female patients have a greater risk of death compared to males. The probability risk was given the equation $R = 1 - e^{-H(t)}$ where $e^{-H(t)}$ is the survival function, $H(t)$ the cumulative hazard function which was created using Cox-Regression.

1. Introduction

In 2012, the National Kidney and Transplant Institute (NKTI) cited kidney failure as the 9th leading cause of death among Filipinos. Most people know that a major function of the kidneys is to remove waste products and excess fluid from the body. These waste products and excess fluid are removed through the urine. The production of urine involves highly complex steps of excretion and re-absorption. This process is necessary to maintain a stable balance of body chemicals.

A diagnosis of End Stage Renal Disease (ESRD) means that you are in the final stage of kidney disease and your kidneys are not functioning well enough to meet the needs of daily life. It also used to be an automatic death sentence; however, advancements in treatments now allow patients to live much longer than ever before. Without dialysis or a kidney transplant, death will occur. The outcome of the treatment depends on the patient.

Having kidney disease increases your chances of also having cardiovascular disease, heart attacks, and strokes. Also, the body can hold in too much fluid which could lead to swelling in your arm and legs, high blood pressure, or fluid in your lungs called pulmonary edema. To analyse the random cause of the morbidity among Filipino people, the researchers generated a mathematical survival model that will predict the absolute risk of death with ESRD patients. This study aimed to determine what variables affect the probability of survival of Chronic Kidney Disease patients diagnosed with ESRD and to assess the effectiveness of the developed survival model to provide reasonably accurate failure analysis and failure forecasts of the conditional status of patient.
2. Statement of the Problem
This study aimed to undertake a survival analysis of ESRD patients. Specifically, it attempted to find answers to the following questions:
1) What is the risk profile of ESRD patients diagnosed with Pulmonary Congestion and Cardiovascular Disease that leads to death?
2) What are the survival functions of patients with ESRD, ESRD with Cardiovascular Disease, Pulmonary Congestion, and Both Cardiovascular Disease and Pulmonary Congestion according to age and gender group using Kaplan–Meier Product Limit Estimates and Weibull Distribution?
3) What are the hazard functions of patients with ESRD, ESRD with Cardiovascular Disease, Pulmonary Congestion, and Both Cardiovascular Disease and Pulmonary Congestion according to age and gender group using Kaplan–Meier Product Limit Estimates and Weibull Distribution?
4) Is there a significant difference between Kaplan-Meier Limit Estimates and Weibull Distribution according to age and gender group when generating survival and hazard rates?
5) What are the probability risks of death of patients with ESRD, ESRD with Cardiovascular Disease, Pulmonary Congestion, and Both Cardiovascular Disease and Pulmonary Congestion using Cox-Regression and Weibull Distribution?

3. Methodology
This section describes the data and methods used in this study to come up with the development of a survival model and estimate the probability of surviving all causes of death for a specified time interval calculated from the cohort of ESRD cases.

3.1. Data Gathering Procedure
In this study, all of the data needed by the researchers were obtained from the records of V. L. Makabali Memorial Hospital. The data collected are from the total number of Chronic Kidney Disease patients diagnosed with ESRD from 2008 to 2013. These consist of the age, gender, and final diagnosis of ESRD patients that led them to death.

To compute the risk of death on CKD patients diagnosed with ESRD, the status of each patient are needed to be observed. The status of the patients is 1 if the person failed, meaning the person died and 0 if the person was censored or alive over the given time. Thereafter, the data must be organized to fit in statistical packages like Number Cruncher Statistical System (NCSS), Statistical Analysis and Graphics and Statistical Package for the Social Sciences (SPSS).

3.2. Kaplan-Meier Method
The Kaplan–Meier (KM) method is a nonparametric survival analysis that is used to estimate the two survival probabilities curves dealing with differing survival times especially when not all of the subjects continue in the study. The assumptions of KM method are:
1. The subjects are a random sample from the population of interest, so that they are independent of each other.
2. Censored individuals have the same prospect of survival as those who continue to be followed. This cannot be tested for and can lead to a bias that artificially reduces $S$.
3. The Kaplan-Meier survival estimate is a step function that changes at every distinct survival time, but does not change at censoring times (unless a survival time happens to be tied to a censoring time).
4. The event studied (e.g. death) happens at the specified time. Late recording of the event studied will cause artificial inflation of $S$.

3.3. Log-Rank Test
The log–rank test is a large-sample chi-square test that uses as its test criterion a statistic that provides an overall comparison of the KM curves being compared. The main objective is to compare two
survival curves by treatment group. Here we consider two groups. The hypotheses of interest are stated as follows:

\[ H_0: S_1(t) = S_2(t) \]
\[ H_1: S_1(t) \neq S_2(t) \]

where the null hypothesis \( H_0 \) states that the survival curves are identical in the two populations and the alternative hypothesis \( H_1 \) states otherwise.

3.4. Weibull Distribution

The Weibull distribution is worldwide used to model life data. The distribution can handle increasing, decreasing or constant failure-rates and can be created for data with and without suspensions (non-failures). The Weibull distribution is flexible and fits to a wide range of data, including Normal distributed data. Only log–normal data does not fit in the Weibull distribution and needs separate analyses. For creating the plot you need to record the time to failure that can be expressed in mileage, cycles, minutes, strength, stiffness or similar continuous parameters. The Weibull cumulative distribution function (CDF), has the explicit equation:

\[ F(t) = 1 - e^{-\left(t/\eta\right)^\beta} \]

where:
- \( F(t) \) := probability of failure at time \( t \);
- \( t \) := time, cycle, miles or any appropriate parameters;
- \( \eta \) := characteristic life or scale parameter; and
- \( \beta \) := slope or shape of the parameter.

3.4.1. Anderson–Darling Test. The Anderson-Darling test is used to test if a sample of data came from a population with a specific distribution. The Anderson-Darling test is an alternative to the chi-square and Kolmogorov-Smirnov goodness of fit tests. The hypotheses of interest are stated as follows:

\[ H_0: \text{The data follows Weibull distribution.} \]
\[ H_1: \text{The data do not follows Weibull distribution.} \]

3.5. Cox Regression

It is nonparametric in the sense that it involves an unspecified arbitrary baseline hazard function. It is parametric because it assumes a parametric form for the covariates. The baseline hazard function is scaled by a function of the model’s (time–independent) covariates to give a general hazard function. It is usually written in terms of the hazard model formula,

\[ H(t, X) = h_0(t) e^{\sum_{i=1}^{n} \beta_i X_i} \]

An important feature of this formula, which concerns the proportional hazards assumption, is that the baseline hazard is a function of \( t \), but does not involve the \( X \)'s. In contrast, the exponential expression shown here, involves the \( X \)'s, but does not involve \( t \). The \( X \)'s here is called time–independent \( X \)'s.

3.5.1. Graphical Method. There are two types of graphical techniques to check the proportional hazard assumptions. The most common technique is by comparing the estimated \( \ln(-\ln(S(t))) \). If the two survival curves do not intersect and are parallel, it clearly provides evidence against the assumptions. Another graphical approach is by comparing observed with predicted survival curves. If the two survival curves are close, then the proportional hazard assumption is plausible.

3.5.2. Test of Goodness Fit. This approach provides a Chi–Square statistics where the computed values for each variable rely on p–values. This p–value is used for evaluating the PH assumption for each variable. If the p–values are large then PH is satisfied, whereas small p–values provide violation of PH assumption.

3.5.3. Probability Risks. The risk of Coronary Heart Disease is given by
\( R(X, \text{Death}) = 1 - e^{-H(t,X)} \),

where \( e^{-H(t,X)} \) is the survival function in terms of its cumulative hazard rate function \( H(t,X) \). Since there were more covariates involved in the study, the researchers used the Cox Regression to estimate the hazard rate of the event on ESRD patients. The formula is given by

\[ H(t,X) = h_0(t) e^{\sum_{i=1}^{n} \beta_i X_i} \]

where \( h_0(t) \) is the cumulative baseline hazard function. This term depends on time but not the covariates, \( X_i \) is possibly time–independent covariables, and \( \beta \) is the regression coefficients.

4. Results and Discussion

The study examined the data collected for trends and differences between independent variables such as ESRD, Pulmonary Congestion, Cardiovascular Disease, age and gender. This section will also interpret the data for further statistical analysis regarding the effects of chosen predictors on death.

4.1. Risk Profile of V.L. Makabali Memorial Hospital Patients

Table below presents the risk profile of 289 V.L. Makabali Memorial Hospital patients with End Stage Renal Disease (ESRD). Of these 97 have complications such as Pulmonary Congestion and Cardiovascular Disease that leads to death.

<table>
<thead>
<tr>
<th>Category of Patients according to Diagnosis</th>
<th>Number of Individuals</th>
<th>Male</th>
<th>Female</th>
<th>Greater than 52</th>
<th>Less than or equal to 52</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESRD</td>
<td>148</td>
<td>78 (53.79%)</td>
<td>70 (48.61%)</td>
<td>97 (50.78%)</td>
<td>51 (52.04%)</td>
</tr>
<tr>
<td>ESRD with Pulmonary Congestion</td>
<td>45</td>
<td>21 (44.88%)</td>
<td>24 (53.33%)</td>
<td>28 (41.17%)</td>
<td>17 (41.17%)</td>
</tr>
<tr>
<td>ESRD With Cardiovascular Disease</td>
<td>54</td>
<td>22 (41.38%)</td>
<td>32 (59.26%)</td>
<td>38 (57.66%)</td>
<td>16 (29.63%)</td>
</tr>
<tr>
<td>ESRD with both Pulmonary congestion and cardiovascular disease</td>
<td>42</td>
<td>24 (56.81%)</td>
<td>18 (42.22%)</td>
<td>28 (42.22%)</td>
<td>14 (33.33%)</td>
</tr>
<tr>
<td>TOTAL</td>
<td>289</td>
<td>145</td>
<td>144</td>
<td>191</td>
<td>98</td>
</tr>
<tr>
<td>ESRD that leads to death</td>
<td>6</td>
<td>2 (4.76%)</td>
<td>4 (7.27%)</td>
<td>3  (4.68%)</td>
<td>3  (9.09%)</td>
</tr>
<tr>
<td>ESRD with Pulmonary Congestion that leads to death</td>
<td>26</td>
<td>11 (42.31%)</td>
<td>15 (57.69%)</td>
<td>15 (57.69%)</td>
<td>11 (40.74%)</td>
</tr>
<tr>
<td>ESRD With Cardiovascular Disease that leads to death</td>
<td>37</td>
<td>15 (40.54%)</td>
<td>22 (59.46%)</td>
<td>28 (59.46%)</td>
<td>9 (24.32%)</td>
</tr>
<tr>
<td>ESRD with both Pulmonary congestion and cardiovascular disease that leads to death</td>
<td>28</td>
<td>14 (50.00%)</td>
<td>14 (50.00%)</td>
<td>18 (64.29%)</td>
<td>10 (35.71%)</td>
</tr>
<tr>
<td>TOTAL</td>
<td>97</td>
<td>42</td>
<td>55</td>
<td>64</td>
<td>33</td>
</tr>
</tbody>
</table>

Table 1 also shows separate gender data for males and females and for age greater than 52 and less than or equal 52 (expressed in counts and %). Thus, for 148 ESRD patients without complications, 78 (or 53.79%) are males while 70 (or 48.61%) are females. In addition, 97 (or 50.78%) are with ages greater than 52 while 51 (or 52.04%) with ages less than or equal 52.

It also shows that there are more female ESRD patients with cardiovascular disease (22.22%) than patients with pulmonary congestion (14.48%); and there are 16.55% of male ESRD patients with both complications. While for the female ESRD patients, 48.61% have no complications. There are more
female ESRD patients affected by cardiovascular disease (22.22%) than those with pulmonary congestion (16.67%); and there are only 12.5% of female ESRD patients with both complications.

In patients of ages greater than 52, there are 50.78% with End Stage Renal Disease only, 14.66% with Pulmonary disease, 19.90% with Cardiovascular disease, and 14.66% with both complications while patients with ages less than or equal to 52, 52.04% have end stage renal disease only. 17.34% of the patients have pulmonary congestion and 16.32% have cardiovascular disease, while 14.28% of the patients have both complications.

From the group of male ESRD patients, 4.76% had died with only ESRD while 26.19% died having ESRD with pulmonary congestion. There are 27.27% of male ESRD patients who died having cardiovascular disease while there are 33.33% of male patients died having both complications. From the female group, 7.27% died while having ESRD. There are more female ESRD patients that died with cardiovascular than patients with pulmonary congestion with a percentage of 27.27% and 40%. From the female group, 33.33% of them died while having both complications; it is the same with the male patients.

On the other hand, for patients with ages greater than 52 years, 4.68% died with only End Stage Renal Disease, 23.44% died with pulmonary congestion, 43.75% died with cardiovascular congestion and 28.13% died with both complications. For patients with ages less than 52 years, 9.09% died with end stage renal disease only while 33.33% died with pulmonary congestion. Moreover 27.27% of patients died with cardiovascular disease and also 30.30% died with both complications.

4.2. Survival functions of patients with ESRD only, ESRD with Cardiovascular Disease only, ESRD with Pulmonary Congestion only, and ESRD with both Cardiovascular Disease and Pulmonary Congestion using Kaplan-Meier Product Estimates and Weibull Distribution.

4.2.1. Age Group.

Figure 1 shows the survival functions according to age of the patients with ESRD, ESRD with Cardiovascular disease, ESRD with Pulmonary Congestion, and ESRD patients with both complications. Kaplan-Meier Estimates is represented by a step function while the Weibull distribution is represented by an exponential function. The vertical axis stands for the probability of survival of patients with respect to age and its complications. The figures show that as the age of the patient increases the probability of its survival decreases.
4.2.2. Gender Group.

![Figure 2. Survival Functions According to Gender Group](image)

Figure 2 shows the survival functions according to gender of the patients with ESRD, ESRD with Cardiovascular disease, ESRD with Pulmonary Congestion, and ESRD patients with both complications. Kaplan-Meier Estimates is represented by a step function while the Weibull distribution is represented by an exponential function. The vertical axis stands for the probability of survival of patients with respect to age and its complications. The probability of survival of male and female patient decreases with respect to age.

4.3. The hazard functions of patients with ESRD, Pulmonary Congestion, and Cardiovascular Disease and with both Pulmonary Congestion and Cardiovascular Diseases using Kaplan-Meier Limit Estimates and Weibull Distribution

4.3.1. Age Group.

![Figure 3. Hazard Functions of Patient According to Age Group](image)
Figure 3 shows the hazard functions according to age group of the patients with ESRD, ESRD with Cardiovascular disease, ESRD with Pulmonary Congestion, and ESRD patients with both complications using Kaplan-Meier Limit and Weibull Distribution. Note that under the age group, it shows that patients whose ages are either greater or less than 52 follow linearity.

4.3.2. Gender Group.

Figure 4 shows the hazard functions according to gender group of the patients with ESRD, ESRD with Cardiovascular disease, ESRD with Pulmonary Congestion, and ESRD patients with both complications using Kaplan-Meier Limit and Weibull Distribution. Note that under the gender group, it shows that patients whose ages are either male or female follow linearity.

4.4. Is there a significant difference between Kaplan-Meier Limit Estimates and Weibull Distribution when generating survival and hazard rates?

4.4.1. Age Group.

Table 2. Significant Difference of the Survival and Hazard Rates under Age Group between Kaplan-Meier Limit Estimates and Weibull Distribution

<table>
<thead>
<tr>
<th>Survival Rates</th>
<th>Computed t-value</th>
<th>p-value</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than or equal to 52 years old</td>
<td>ESRD</td>
<td>2.364</td>
<td>0.142</td>
</tr>
<tr>
<td></td>
<td>ESRD with Cardiovascular Disease</td>
<td>9.040</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>ESRD with Pulmonary Congestion</td>
<td>4.323</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>ESRD with both complication</td>
<td>7.762</td>
<td>0.000</td>
</tr>
<tr>
<td>Greater than 52 years old</td>
<td>ESRD</td>
<td>5.168</td>
<td>0.122</td>
</tr>
<tr>
<td></td>
<td>ESRD with Cardiovascular Disease</td>
<td>8.306</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>ESRD with Pulmonary Congestion</td>
<td>5.098</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>ESRD with both complication</td>
<td>6.796</td>
<td>0.000</td>
</tr>
</tbody>
</table>
Table 2 shows that there is no significant difference between the survival rates generated by Kaplan-Meier Estimates and Weibull Distribution under age groups of patients with ESRD for both survival and hazard rates and also the hazard rates for ESRD with Pulmonary Congestion whose age is greater than 52 years old. And the other combinations are all significant.

4.4.2. Gender Group

Table 3 shows that there is no significant difference between the survival rates generated by Kaplan-Meier Estimates and Weibull Distribution under gender groups of patients with ESRD for both survival and hazard rates and also the survival and hazard rates for ESRD with Pulmonary Congestion female patients. And the other combinations are all significant.

4.5. What are the probability risks of death of patients with ESRD, Pulmonary Congestion, Cardiovascular Disease and with Both Cardiovascular Disease and Pulmonary Congestion using Cox-Regression and Weibull Distribution?

4.5.1. Probability risk of Death on ESRD patients under Cox Regression. The probability risk of Death is identified by the model $R(x, \text{Death}) = 1 - e^{-H(t|x)}$ where $e^{-H(t|x)}$ is the survival probability in
terms of a cumulative hazard rate function $H(t,X)$ which is derived from Cox Proportional Hazard Model and can be verified in the appendices.

Substituting the Cox proportional hazard rate function to the model, the following probability risk functions:

1. ESRD with Pulmonary Congestion
   \[ R(x, \text{Death}) = 1 - e^{-(0.000715e^{0.655t-2.639x_1-0.206x_2})} \]
   where $t =$ age of individual
   \[ x_1 = \begin{cases} 1, & \text{if affected by ESRD} \\ 0, & \text{if affected by ESRD with Pulmonary Congestion} \end{cases} \]
   \[ x_2 = \begin{cases} 1, & \text{if male} \\ 0, & \text{if female} \end{cases} \]

2. ESRD with Cardiovascular Disease
   \[ R(x, \text{Death}) = 1 - e^{-(0.000258e^{0.086t-2.546x_3-0.262x_4})} \]
   where $t =$ age of individual
   \[ x_3 = \begin{cases} 1, & \text{if affected by ESRD} \\ 0, & \text{if affected by ESRD with Cardiovascular Disease} \end{cases} \]
   \[ x_4 = \begin{cases} 1, & \text{if male} \\ 0, & \text{if female} \end{cases} \]

3. ESRD with both Cardiovascular Disease and Pulmonary Congestion
   \[ R(x, \text{Death}) = 1 - e^{-(0.00045e^{0.075t-2.654x_5-0.204x_6})} \]
   where $t =$ age of individual
   \[ x_5 = \begin{cases} 1, & \text{if affected by ESRD} \\ 0, & \text{if affected by ESRD with Cardiovascular Disease and Pulmonary Congestion} \end{cases} \]
   \[ x_6 = \begin{cases} 1, & \text{if male} \\ 0, & \text{if female} \end{cases} \]

4.5.2. Probability Risk of Death on ESRD Patients under Weibull Distribution. The probability risk of Death in Weibull Distribution is identified by the model $R(x, \text{Death}) = 1 - e^{-H(t,X)}$ where $e^{-H(t,X)}$ is the survival probability in terms of a cumulative hazard rate function $H(t,X)$.

Substituting the Weibull Distribution hazard rate function to the model, the researchers’ got the general probability risk functions of patients with Cardiovascular Disease, Pulmonary Congestion and with both Disease according to the gender of the patients. For female the model is
   \[ R(t, \text{Death}) = 1 - e^{-(0.000258e^{0.086t-2.546x_3-0.262x_4})} \]
   And for male the model is
   \[ R(t, \text{Death}) = 1 - e^{-(0.00045e^{0.075t-2.654x_5-0.204x_6})} \]
where $t =$ age of individual.

5. Conclusion
According to the results, the researchers conclude that having End Stage Renal Disease (ESRD) with complications increases the probability of death. In addition, female patients have greater risk of death compared to males and patients with ages greater than 52 years have a higher risk of death compared to patients with ages less than 52. The estimated survival and hazard rates of ESRD patients under age and gender group by Kaplan-Meier Method and Weibull Distribution has no significant difference with p-values greater than 0.05 and also there is no significant difference between the hazard rates of
male ESRD patients with both Cardiovascular Disease and Pulmonary Congestion with p-value equals to 0.08. The developed probability risk of death in ESRD patients generated by Cox regression and Weibull Distribution at certain age will help medical experts to improve the quality of medications and technologies that will lessen the risk of death in ESRD and to patients with or without ESRD as a base of healthy lifestyle living.

6. Recommendation
End-stage renal disease is often preventable with the proper control of blood pressure, blood lipids, and blood glucose, in combination with appropriate medications. For patients who have progressed to the need for dialysis, morbidity and mortality can be reduced and quality of life enhanced through adherence to an appropriate diet and medical regimen, along with regular physical activity. Future researchers may include other factors such as Body Mass Index, blood counts or tests, blood pressure and urine test as independent variables for the risk of death. The researchers suggest the use of more samples from different hospitals that would make the model a general representation for estimating the risk of death.

7. References