SURVIVAL ANALYSIS OF TIME-TO-FIRST PERITONITIS AMONG KIDNEY PATIENTS WHO ARE ON PERITONEAL DIALYSIS AT PIETERSBURG PROVINCIAL HOSPITAL, LIMPOPO PROVINCE, SOUTH AFRICA

by

TSHEPO FRANS MAJA

DISSERTATION

Submitted in fulfillment of the requirements for the degree of

MASTER OF SCIENCE

in

STATISTICS

in the

FACULTY OF SCIENCE AND AGRICULTURE

(School of Mathematical and Computer Sciences)

at the

UNIVERSITY OF LIMPOPO

SUPERVISOR: Prof. YG KIFLE

CO-SUPERVISOR: Dr. RA TAMAYO ISLA

30 NOVEMBER 2020

Declaration

I Tshepo Frans Maja, hereby declare that this dissertation submitted to the University of Limpopo for the degree of Master of Science in statistics is my original work and has never been submitted to any other university or institution and all materials used therein has been respectfully acknowledged.

Sign.....

Tshepo Frans Maja

Date: 30 NOVEMBER 2020

Abstract

Peritoneal Dialysis (PD) is a process of replacing kidney function which cleans waste from the blood and remove extra fluid from the body. In most cases, the process of PD is slowed down by a peritoneal membrane infection called peritonitis. Despite recent advancements in treatments and prevention, peritonitis still remains the leading complication which results in high morbidity and technique failure among PD patients. Using a prospective peritonitis dataset of 159 kidney patients who were on PD from 2008 to 2015 in Pietersburg Provincial Hospital, the aim of this study was to identify potential social, demographic and biological risk factors that contribute to the first episode of peritonitis. Both semi-parametric (Cox PH) and parametric (Accelerated Failure Time: Weibull, exponential, log-logistic, and gamma) survival models were fitted to the peritonitis dataset. Akaike Information Criterion (AIC) was applied to select models which best fit to the peritonitis data. Accordingly, log-logistic Accelerated Failure Time (AFT) model was found to be a working model that best fit to the data. A total of 96 (60.38%) peritonitis cases were recorded over the follow-up period with majority of peritonitis infection coming from females (65.4%) and rural dwellers (65.7%) with (62.6%) of black Africans showing higher risk of developing peritonitis. The multivariate log-logistic AFT model revealed that availability of water (p-value=0.018), electricity (p-value=0.018), dwelling (p-value=0.008), haemoglobin status (p-value=0.002) and duration on PD (p-value=0.001) are significant risk factors for the development of peritonitis. Therefore, patients with no water and electricity, coming from rural background with low level of haemoglobin and shorter duration on PD are associ-

ated with high risk or hazard of developing peritonitis for the first time.

Dedication

This dissertation is dedicated to my guiding light, Almighty God, my late parents, Mr and Mrs Maja, who taught me the importance of education and hard work. I would also like to dedicate this dissertation to my uncle, Mr Madimetja Abram Maja, who played a crucial role through out my schooling program. I am eternally grateful and appreciative for his continuous support through the tough times of my educational journey.

Acknowledgements

This dissertation could not have been written without my supervisor Prof. YG Kifle and co-supervisor Dr RAT Isla who did not only serve as my advisors but also challenged and encouraged me throughout my dissertation program. Additionally, this dissertation could not have been a success without their constant support, advice and patience. I am deeply indebted to Prof. YG Kifle who opened my eyes for survival analysis modelling techniques and guided me throughout. He always created time for me despite his busy schedule. I appreciate the way in which he explained very difficult concepts in very simple ways. Thank you for your support, words cannot express my gratitude. I am extremely grateful to the Department of Statistics and Operations Research at the University of Limpopo for giving me the opportunity to pursue my dream. Last but not least, I would like to thank my friends and family, for their support and encouragement.

Contents

List of Figures

List of Tables

List of Acronyms

PD	Peritoneal Dialysis
AIC	Akaike Information Criterion
CKD	Chronic Kidney Disease
CAPD	Continuous Ambulatory Peritoneal Dialysis
CCPD	Continuous Cycling Peritoneal Dialysis
ESRD	End-Stage Renal Disease
HD	Haemodialysis
RRT	Renal Replacement Therapy
ISPD	International Society of Peritoneal Dialysis
PKDC	Polokwane Kidney Dialysis Center
KM	Kaplan Meier
AFT	Accelerated Failure Model
PH	Proportional Hazard
RRF	Residual Renal Function
SAS	Statistical Analysis System
HR	Hazard Ratio
TR	Time Ratio
PH	Proportional Hazard
AFT	Accelerated Failure Time
LVH	Left Ventricular Hypertrophy
PHT	Pulmonary Hypertension
PTH	Parathyroid Hormone

Chapter 1

Introduction and background

1.1 Peritoneal dialysis

Peritoneal dialysis (PD) is a treatment for people who have kidney failure. Kidney failure is stage five of chronic kidney disease which occurs when the kidneys of a patient are malfunctioning. The waste and extra fluid can build up in the patient's blood and make them sick. PD removes waste product from the blood through the peritoneal membrane (the lining of the belly). The peritoneal membrane is a membrane that surrounds all of our organs. The membrane acts as a natural filter which filters out waste products and extra fluid in the person's blood pass through the cleansing fluid (dialysate) and at the same time holds back essential things such as nutrients and red blood cells needed by the body. In PD, a catheter is placed in the abdomen and the dialysate is instilled into the abdomen and left for a number of hours. This time is called the patient's dwell time. After the prescribed amount of time by the doctor, the dialysate is drained out of the abdomen. The in-and-out process is repeated a number of times during the day, using a fresh

dialysate. Patients have different peritoneal characteristics, and require individualised PD regimes. There are two types of home PD, namely Continuous Ambulatory Peritoneal Dialysis (CAPD) and Continuous Cycling Peritoneal Dialysis (CCPD).

1.1.1 Continuous ambulatory peritoneal dialysis

A patient undergoing CAPD treatment connects a fill-and-drain bag to the catheter, performs the exchange and disconnect the tubing from the catheter three to four times a day. It is a manual method of doing the exchanges. Therefore, precautions must be taken to ensure the sterile connection to the catheter. However, the method is complicated by peritonitis rate, because the technique involves spiking class bottles of PD up to almost four times per day. This can lead to the production of PD fluid inside the plastic bag which may remain attached to the patient for the whole dwell time prescribed by the doctor.

1.1.2 Continuous cycling peritoneal dialysis

CCPD, which is also known as automated peritoneal dialysis, is a method of doing the exchanges with the machine. A machine called a cycler does the exchange automatically while a patient is asleep. The patient connects his catheter to the cycler at midday, which automatically drains-and-fill abdomen once. This process is repeated four to five times at bedtime whiles the patient sleeps for eight to ten hours.

1.2 The infection in peritoneal dialysis

Peritonitis, which is an infection of the peritoneal membrane, remains a major complication of PD, because it can lead to catheter loss, hospitalisation, and technique

failure (Kerschbaum et al., 2012). Peritonitis rates vary across the world, but reasons for these are unclear.

1.3 Background of the study

PD is an established treatment for end-stage renal disease (ESRD). This is made possible by outstanding progress in science and medicine that were conducted in the eighteenth and nineteenth centuries (Feng et al., 2016; Negoi and Nolph, 2009). According to Brown (2012), PD was used before haemodialysis (HD) in 1923, but the difficulties as a result of infective complications and gaining access to the peritoneal cavity effectively suggested further improvement of this technique. Originally, PD was tested in women with renal failure in 1923, although an initial experience with intermittent PD was discouraging (Mehrotra and Boeschoten, 2009). It was discovered that PD is not a good or appropriate renal replacement therapy (RRT) for patients with ESRD. This was due to initial unsuccessful attempts at designing indwelling catheters because of infections complications (Mehrotra and Boeschoten, 2009).

In the late 1960s intermittent PD was practiced worldwide and it became a safe and standardized procedure (Oreopoulos and Thodis, 2010). Patients would come to the hospital once or twice a week and be put on PD with a new catheter inserted into their abdomen or peritoneal cavity (Brown, 2012). There was an improvement in the treatment of PD with the introduction of indwelling silicon rubber catheter with cuffs designed by Tenckhoff (Oreopoulos and Thodis, 2010). This new improved catheter was able to reduce risk of infections and therefore it became a long term treatment option (Oreopoulos and Thodis, 2010). To this date, the original Tenckhoff catheter still remains the main used catheter for chronic peritoneal access (Oreopoulos and Thodis, 2010).

According to Nayak et al. (2009), PD should be the preferred modality of RRT for

patients with ESRD in developing countries due to its inherent advantages. The motive underlying this treatment, is that PD is less costly, home based therapy even in rural areas, enhances quality of life and preserve residual renal function (RRF) (Nayak et al., 2009). Blake (2001) asserted that PD may have significant advantages from variety of social, medical and economic perspectives over HD and these advantages may be witnessed at an early stages or years. Moreover, Nayak et al. (2009) highlighted that in Africa, RRT remains a major challenge for renal communities because of lack of motivated health care teams, infrastructure and costs. In addition, they indicated that successful countries such as South Africa and those from North Africa (Algeria, Libya, Morocco and Tunisia) have developed a well established PD programs. Although PD is faced with several setbacks, they demonstrated that developing countries still hold the key of PD as therapy in the future (Nayak et al., 2009).

In 1976 a new option of PD technique meant to treat chronic renal failure called CAPD was developed by Popovich and colleagues (Moncrief and Popovich, 1979). According to Stason et al. (1985), the advent of CAPD dramatically changed the treatment options available for ESRD patients and indicated that the use of CAPD has increased rapidly.

Winterbottom (2015) stated that the Tenckhoff catheter should be inserted in the abdomen through a minor operation, which acts as a channel for the fluid exchange. The technique consists of four exchanges daily of at least two litre to produce, with total dialysate and ultrafiltration of about 10 litre per day. Gokal and Mallick (1999) affirmed that CAPD does not require specific and complex equipment and it is at least 25% cheaper than in-hospital HD, with hospital cost included. Next, is the development of CCPD, which employs a cycler that controls the volume, dwell time, filling and drainage of fluid when a patient is asleep. According to Fourtounas (2011), CCPD is growing fast as a PD modality all over the world. This is due to its ability to allow a patient to undergo his/her therapy at night. It allows the patient and his/her helper to be free during the day with short-dwell cy-

cles which runs in-and-out of the peritoneal cavity by cycler machine. Moreover, it offers an advantage for fast transport as the dwell time enables the patients to have an increased ultrafiltration per exchange.

Despite recent advances in PD systems or regimes, peritonitis remains the leading complication and the most significant biological problem in patients who are on PD. According to Taha et al. (2017), about 18% of all infections in PD kidney patients are as a result of peritonitis. In South Africa, peritonitis has become one of the growing concerns to most health care departments and kidney patients in recent times. According to Piraino (1998), peritonitis remains a major problem for patients on PD. This is the main primary cause of high morbidity, technique failure, catheter loss, possible permanent membrane damage, transfer to HD, and in some circumstances can be fatal (Mustafa et al., 2015; Martin et al., 2011; Kelecş et al., 2010). The usual typical biological symptoms of peritonitis is vomiting, abdominal pain, fever or if there is a cloudy effluent observed or if an effluent cell count with white blood cells of at least $100/\mu L$, with more than 50 % polymorphonuclear neutrophilic cells (Isla et al., 2014). The abdominal pain can range from extreme severe to non existent, whereby, the absence of the abdominal pain in inexperienced patients, may lead them to ignore the initial cloudy effluent, leading to delay in the subsequent treatment.

Many researchers have come out with causes and complications of peritonitis in South Africa and elsewhere. Although reports are inconsistent, some of the factors associated with increased peritonitis risk have been documented. For, instance, Isla et al. (2014); Feng et al. (2016), discovered a link between the causes of peritonitis in PD and other risk factors such as age, gender, race, employment status, body mass index (BMI) [kg/m^2], level of education, tap water present, electricity present, marital status, level of income per month, hemoglobin, duration on PD, distance from PD center (km) and causes of ESRD such as diabetes and hypertension. However, according to Nieto-Ríos et al. (2014), the incidence of PD-related peritonitis varies among different dialysis center and the period of time. There-

fore, among these factors, some could be possible contributors of the occurrence of time-to-first peritonitis in Polokwane Kidney Dialysis Center (PKDC) in Limpopo province, South Africa. In addition, there are a number of micro-organisms responsible for peritonitis depending on the number of covariates. Furthermore, due to different risk factors such as poverty, education levels, climate change, lack of water and electricity, developing countries are the worst affected with this infection (Ikabu et al., 2016b). Peritonitis may complicate the utilisation of PD due to its severity in term of endurance and may lead to the failure of the peritoneal membrane, which may force the patient to transfer from PD to another treatment such as HD.

Figueiredo et al. (2013), discussed that the International Society of Peritoneal Dialysis (ISPD) in 2005 recommended that monitoring of peritonitis in the dialysis units should be performed by calculating the rate of episode per year at risk. Moreover, they highlighted that the 2010 ISPD guidelines introduced the recommendation to establish the median peritonitis rate on PD programs, wherein the peritonitis rate is calculated by episode per year by patient. Previously, ISPD guidelines focused primarily on protocols of treating peritonitis, but not the prevention measures (Bender et al., 2006). Nevertheless, the current guidelines focus more attention on preventive measures. According to Bender et al. (2006), reducing of peritonitis infection should be the top priority of every PD program, which clearly shows that the appropriate treatment is clearly important in the event of the infection.

The role of peritonitis as an independent risk factor for technique failure and mortality in PD kidney patients has been established for several decades (Feng et al., 2016). A study conducted in Colombia by Nieto-Ríos et al. (2014), found the incidence of peritonitis to be 0.84 episode per patient. Feng et al. (2016) reported that the first peritonitis episode can change peritoneal membrane function.

It is however surprising that in spite of the numerous factors identified by researchers in South Africa and elsewhere as the contributors of peritonitis in various

dialysis centers and its consequences among kidney patients on PD, few studies have been conducted in Limpopo province. According to Faber (2003), the risk factors for peritonitis are identifiable and modifiable and requires focus intervention.

1.3.1 Background of Limpopo province

Limpopo province is one of the nine provinces of the Republic of South Africa situated in the north eastern part of the country. The province covers an area of 123 910 km^2 with estimated population of 5.3 million (Gafar, 2013). The province is divided into five administrative districts namely Sekhukhune, Vhembe, Capricon, Waterberg and Mopani, and it represents almost 12% of SA's population (Igumbor and Laubscher, 2003). The province is 89.3% rural, with 71.8% unemployment rate. The province consists of several ethics groups distinguished by race, culture and language, with about 97% of the total population being black African, 2.6% white, 0.2% Indian and 0.2% coloured. The Sepedi language constitute the largest spoken language, with about nearly 57% of the population of the province, followed by Tsonga language which comprise 23% and Venda speakers who constitute 12%. The Afrikaans and English speaking whites make up 2.6% and less than 0.5%, respectively (Gafar, 2013). The province is ranked sixth of all the provinces in South Africa in terms of total income. In per capita income terms, however, the province is the poorest (Pauw et al., 2005). The province is unique in that, it serves as a link between South Africa and Zimbabwe, Botswana, and Mozambique which are afield in the sub-Saharan Africa. The capital city of the province is Polokwane, where this study is being conducted and is situated in the Capricon District, which is the most central of the five districts.

1.3.2 Dialysis center and peritonitis in Polokwane (Limpopo province)

The PKDC, part of the Pietersburg Provincial Hospital is the only dialysis unit in the public sector in the province and was officially opened in 2007 (Isla et al., 2014). The center specifically offered services on PD and HD for kidney patients (Isla et al., 2014). According to Isla et al. (2014), South African government policy permitted only 50 PD patients to undergo the RRT program at PKDC. However, the dialysis unit collaborated with Fresenius Medical Care, hence, the collaboration is now public-private partnership. As a result of this partnership, some capital was invested into the center and this made it possible to admit more kidney patients (Isla et al., 2014). This collaboration meant that locally PD fluids and imported PD fluids are equally utilised in the dialysis unit.

Most patients undergoing PD at the center usually start with 4 exchanges of two liter bags (1.5 % solution) per day. The prescription for patients who are volume overloaded may change from time to time to permit for better exchanges and removal of the fluid. Currently there is no capacity for renal transplant in the province and all patients are cared for by specialist general physician in the dialysis unit. Furthermore, according to Isla et al. (2014), there is no nephrologist in the province.

Isla et al. (2014) further reported that, the ISPD guidelines are always used to treat peritonitis and the data on peritonitis in the province were only available from January 2008. Moreover, the report asserted that the peritonitis rate was computed as the number of infections by organism for a time period divided by dialysis year time at risk and it is expressed as an episode per year. Apart from Isla et al. (2014); Ikabu et al. (2016a); Isla et al. (2016); Raaijmakers et al. (2010), the literature is very scarce in South Africa and there are virtually no studies that reported well documented factors that contribute to the development of time-to-first peritonitis.

1.4 Study setup and data structure

This is a prospective study review of PKDC based on the records of all the kidney patients who were on PD between 2008 and 2015. The study's setting is Polokwane in the Limpopo province, South Africa. This center provides health services required by kidney patients to undergo dialysis. These services include catheter, dialysis solution and the training staff which teaches the patients everything they need to know about PD, for instance, how to perform the exchange and cleaning of the catheter.

All kidney patients who initiated PD at the center were registered and relevant social, demographic and biological factors were recorded during the follow-up period. These factors include the present of tap water at home, present of electricity, employment, dwelling, seasons, catheter removal, race, level of education, house type, dialysate, unuric at baseline, sex, BMI, cholesterol, distance from hospital (km), age at baseline, weight, serum albumin, Hb, glomerular filtration rate (eGFR) and causative organisms.

The study population consist of 159 patients' records and were evaluated on a monthly basis. The inclusion criteria is any patient who have been diagnosed with kidney failure and commenced PD between 2008 and 2015 at the center. At each follow-up period, clinical assessments, vital symptoms and biological tests were conducted by the staff team, and details of peritonitis event were also recorded.

To manage the data collection process, a list of patients on record every year was recorded into Microsoft Excel software with all aforementioned factors. This allowed the data to be transferred to other statistical packages such as SAS for analysis. However, patients who had started PD from other provinces and later joined the PKDC after moving to Limpopo province where excluded in the study.

Table 1.1: Time-to-first peritonitis data set for 159 kidney patients who were on PD at Pietersburg Provincial hospital

ID	Time-to-fist peritonitis in months	Status	Age at baseline	Sex	Race	BMI	Hb	water	electricity	dwelling
1	13	0	34	Male	Black	32.6	9.1	Yes	Yes	Rural
2	6	1	37	Female	Black	33.6	8.4	Yes	Yes	Rural
3	30	0	34	Male	Black	21	10.8	Yes	Yes	Urban
.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.
157	21	0	26	Male	Black	28.9	11.6	Yes	Yes	Rural
158	2	1	15	Female	Black	.	5.6	Yes	Yes	Rural
159	22	0	41	Male	Black	22.2	11	Yes	Yes	Rural

The first column contains patient identification number, the second column gives time-to-first peritonitis (in months),the third column gives the censoring status taking value one (status = 1) if the patients have developed time-to-first peritonitis and value zero (status = 0) otherwise. The fourth and fifth columns gives the age at baseline and sex of the patient, respectively. The sixth column gives the race of the patient. The seventh and eighth columns gives the BMI and Hb at baseline for each patient. The ninth and tenth column gives the availability of water and electricity at home for each patient and the eleventh column gives the dwelling to which the patient comes from.

1.5 Statement of the problem

Peritonitis remains a major problem in patients who are on PD (Piraino, 1998). It is associated with high morbidity, technique failure, catheter loss, possible permanent membrane damage, transfer to HD, and in some circumstances it can be fatal (Mustafa et al., 2015; Martin et al., 2011; Kelecş et al., 2010). It usually occurs due to contamination of a bacteria from the skin or the environment of the inserted catheter (Travar et al., 2015). Peritonitis continues to be an issue although there have been a decrease in the incidence rate with recent improvements in PD treatment and advanced technologies (Kelecş et al., 2010). This infection alone contributes 2% to 3% of all mortality and 15% of the technique failure in PD (Barone et al., 2011).

With respect to the modelling aspect, most medical researchers and data analysts utilised logistic regression models to analyse data related to peritonitis and the results of these techniques have contributed considerably to the current knowledge of risk factors associated with peritonitis. However, this modelling approach does

not utilise the complete information available of time-to-event (survival) data. It only takes into account the presence or absence of the infection and ignores the duration required for the infection to occur during the follow-up period. Hence, in this study, we propose the use of various survival data analysis techniques to model time-to-first peritonitis as a function of a number of potential factors available in the data set.

In South Africa and across the world, due to multi-cultural diversity of the dialysis centers, conflicting results related to various factors associated with peritonitis have been reported. However, most of these studies did not utilise time-dependent methods such as survival analysis techniques, therefore leading to diverse conclusions. In PD patients, Chern et al. (2013) indicated that the peritonitis should not simply be determined as the number of peritonitis cases but should also be calculated as a time-to-event function.

Moreover, social and biological factors on the development of peritonitis may cause variation amongst individuals and dialysis centers. And yet, very few studies have been conducted on social, demographic and biological factors leading to the development of first peritonitis episode in Limpopo province. Therefore, the study attempts to fill the gap by identifying social, demographic and biological factors which could lead to the development of first peritonitis episode, which may eventually cause complications in PD using survival analysis techniques.

1.6 Rationale for the study

The motivation behind this study is to use a complete information available in the data set to address potential social, demographic and biological factors contributing to peritonitis in the study area. Hence, application of survival analysis techniques will allow the use of all available information effectively, including the censored observations. Survival models, in contrast with ordinary regression models,

incorporates information from both censored and uncensored patients in estimating model parameters.

In the context of survival analysis of peritonitis patients, we will only take into account the first episode of this infection. In most clinical studies, where patients may develop more than one event in the follow-up period such as peritonitis, the first event is often the primary interest (Zhaojin, 2012). This can limit or reduce a reoccurrence of peritonitis which may occur if the first event has occurred (Zhaojin, 2012). Therefore, the primary motive of analysing the first occurrence of peritonitis is to prolong, if not eliminate, the reoccurrence of the disease by identifying contributing factors from the very beginning.

A number of studies conducted so far on risk factors associated with peritonitis were based on retrospective data (Isla et al., 2014; Han et al., 2007). This current study will be based on prospectively collected data in Pietersburg Provincial Hospital. There has not been much statistical research in the field of peritonitis and PD in South Africa using survival analysis techniques, leading to limited information on potential risk factors contributing to peritonitis.

1.7 Aim of the study

The aim of the study is to analyse the survival of time-to-first peritonitis among kidney patients on dialysis at Pietersburg Provincial Hospital.

1.8 Objectives of the study

The objectives of the study are to:

- Apply various survival analysis techniques to model time-to-first peritonitis episode.

- Compare various survival analysis techniques used to model time-to-first peritonitis episode.
- Identify PD patients who are at high risk for first peritonitis infection.
- Identify potential social-economic, demographic and biological factors that significantly contribute to the risk of peritonitis in the study area.

1.9 Research questions

The study attempts to answer the following questions:

- Which fraction of patients between rural and urban dwellers will survive past a certain time?
- Which social, demographic and biological factors will influence the development of peritonitis in kidney patients who are on PD at PKDC?

1.10 Methodology of the study

The data combined both categorical and quantitative data for the analysis. Secondary data was obtained from PKDC between the period of 2008 to 2015, inclusive. Survival analysis method was used for data analysis. SAS statistical package was used for both data management and analysis purposes.

1.10.1 Survival analysis application

Survival data is collected for the investigation of time-to-event outcome. The event could be death, machine failure, occurrence of the disease, among others. Survival analysis is a statistical method which focuses on the occurrence and duration of events (Rooney, 2015). It examines and models the time it takes for an event to

occur. Typically, it examines the relationship of the survival distributions to factors. One of the important step in survival analysis is to define the event of interest. Accordingly, the study will focus on the occurrence of first peritonitis as an event of interest in kidney patients who are on PD at Pietersburg Provincial Hospital.

Although there are well known methods of estimating unconditional survival distributions, a key characteristic that differentiate survival analysis from other areas of statistics is the fact that survival data are censored. There are three types of censoring, namely right, left and interval censoring. In this study we will focus on the right censoring where the event of interest (peritonitis) is known to have occurred only after a certain time point. Therefore, patients in the study will be censored for the following reasons: firstly, at the end of the follow-up period, the peritonitis will probably not have occurred in all patients. This is because, we do not know when or whether such patients will experience peritonitis, but we know that they were peritonitis free by the end of the observation period. Secondly, the study will censor patients who were lost to follow-up during the study or may have experienced a competing event such as death as a result of which further follow-up is impossible. Thirdly, the transfer to HD for reasons other than peritonitis and finally the transfer of patients from PKDC to other hospitals.

Kaplan-Meier (KM) estimators and log rank test, which are non-parametric methods and Cox proportional hazards (PH) model which is a semi-parametric approach will be utilised in this study. KM estimator and log-rank test will be used to estimate the survival probabilities and compare survival between groups, respectively. Cox PH model will be used to model the effect of various factors on time-to-first peritonitis. At last, accelerated failure time (AFT) parametric modelling approach with exponential, Weibull, log-logistic and gamma distributions will be applied to measure the direct effect of the factors on time-to-first peritonitis.

1.11 Study hypothesis

There is a positive relationship between time-to-first peritonitis and at least one of the risk factors under investigation.

1.12 Scope of the study

This study was conducted at PKDC in Limpopo province, South Africa, using survival analysis techniques and a total number of 159 patients on PD were included in the study. The content covered includes social, demographic and biological factors such as water, electricity, employment, dwelling, season, catheter removal, sex, race, education level, house type, dialysate, BMI, cholesterol, distance from hospital (in km), age at baseline, weight, and Hb which could contribute to first peritonitis episode. The study covers a period from 2008 to 2015, inclusive.

1.13 Significance of the study

The findings of this study could be useful for the Department of Health in designing intervention measures and set appropriate plans to improve the effectiveness of PD and to avoid the infection of peritonitis. The study may also help doctors to identify kidney patients at high risk of peritonitis, allowing for earlier or more frequent counseling of behavioural changes to decrease risk of being infected by peritonitis during the time on PD. The findings could also be helpful for monitoring and evaluating the activities for the government and various agencies concerned. The study will also be a source of reference for other researchers intending to model time-to-first peritonitis under various conditions.

1.14 Summary of the study

Chapter 1 provides an introduction with a background of the study, the problem statement, rationale of the study, the aim of the study, the objective of the study, research questions, methodology of the study, hypothesis of the study, scope of the study, and the significant of the study.

1.15 Organisation of the study

This study contains five chapters but prior to these chapters are the abstracts which gives the summary of the whole research, table of content, list of figures and abbreviations, dedications and acknowledgements.

Chapter 1 of this study contains the introductory information about PD and peritonitis, including the background of peritonitis and Limpopo province, statement of the problem, rationale to the study, aim and objectives of the study, the research questions, methodology of the study, study hypothesis, scope of the study, significance of the study, summary of the chapter and the organisation of the study. Chapter 2 is made up of the relevant literature review, specifically on factors contributing to the occurrence of time-to-first peritonitis in South Africa and elsewhere. Limitations of the study and the statistical techniques that were used to investigate potential risk factors have also been presented. Chapter 3 discusses various statistical survival analysis methods that were used in this study. In Chapter 4, the results and modelling of the data are extensively discussed and analysed whereas Chapter 5 presents the summary of the findings, concluding remarks based on the findings and recommendations. The references are also presented in this study.

Chapter 2

Literature Review

2.1 Introduction

This chapter contains a review of relevant literature on the topic under consideration. This is meant to uncover findings and critical facts which have already been identified by previous studies in and around the causes of peritonitis. It will also present previous studies on peritonitis and various statistical techniques employed. In South Africa, few studies have been conducted in this field, with majority of those consulted obtained from an international perspective. This chapter will guide the direction of this study and help in exploring the unknown.

2.1.1 Definition and classification of peritonitis

Diagnosis and outcomes of peritonitis

The diagnosis of peritonitis episode is said to have occurred if a patient shows the following symptoms, fever, abdominal pain and vomiting, effluent white blood cell count exceeding $100/\mu L$ (after the dwell time of at least two hours), with at least 50% neutrophils (polymorphonuclear neutrophils) which indicates the presence of inflammation with peritonitis being the most likely cause, among others (Isla et al., 2014; Rudnicki et al., 2010; Li et al., 2010; Bernardini et al., 2005; Korbet et al., 1993). However, Taha et al. (2017) indicated that, although patients with peritonitis most often experience severe pains, some peritonitis episodes are associated with mild or even no pain at all.

The outcome of peritonitis was previously described based on the standard definitions in literature and it is a common problem that occurs in patients with ESRD treated with PD. Peritonitis remains the primary cause of high morbidity and technique failure, catheter loss, possible permanent membrane damage, transfer to hemodialysis, and in some circumstances can be fatal, and also one of the major cause of hospitalization (Mustafa et al., 2015; Martin et al., 2011; Kelecş et al., 2010; Bender et al., 2006). Although not more than 4% of peritonitis results in death, peritonitis still remains a contributing factor of death, accounting for almost 16% of death in PD and also a major cause of patients discontinuing PD and switching to HD (Taha et al., 2017). Peritonitis also damages the peritoneal membrane by interfering with its ultrafiltration and the dialysis capacity which may be temporary or permanent (Brown, 2012). According to van Esch et al. (2016), peritonitis was hypothesized to contribute to protein loss and ultrafiltration failure. Indeed, several studies have found high peritonitis rate as a risk factor of ultrafiltration failure, which continues to be the cause of the treatment drop out, and may lead to functional and structural changes which occurs within the peritoneal membrane (Davies et al., 1996).

Time-to first-peritonitis

The first peritonitis episode was defined as the initial event in which the cloudy peritoneal fluid was witnessed, with a leukocyte count of at least 100 $cells/mm^3$, and with more than 50% polymorphonuclear cells (Fan et al., 2014).

Peritonitis rate

There is a wide variation in peritonitis rates among different dialysis centers, countries and period of time (Nieto-Ríos et al., 2014). The reported rates range between 0.06 to 1.66 episode per patient-year and these rates are dominated by single center studies (Cho and Johnson, 2014). According to Cho and Johnson (2014), the sources of the variation in peritonitis rates among different dialysis centers, countries and period of time, relate to the coding bias. Brown (2012), indicated that inconsistence results from different peritonitis rates literature are likely to be caused by study design and differences in the study populations.

The ISPD in 2005, recommended a peritonitis rate of not more than 0.67 episodes per patient-year or 1 episode every 18 months (Rudnicki et al., 2010). However, Rudnicki et al. (2010), also emphasised that centers should strive for lower peritonitis rates such as 1 episode every 41-52 months. The ISPD also published a specific diagnostic criteria for PD-related peritonitis, with an attempt to decrease such variation (Cho and Johnson, 2014). Piraino et al. (2011), indicated that the calculation of peritonitis rates must be standardised and be well defined in any publication on peritonitis. They further stated that, low peritonitis rates are achievable and the infection rates should be monitored by every program for at least quarterly, (Piraino et al., 2011). Bender et al. (2006), also highlighted that if close attention is paid to the causes of peritonitis and the implementation of protocols to reduce the infection, very low rates of peritonitis are possible. Furthermore, peritonitis rates should be calculated for individual organism and be compared to the peritonitis

rates in literature (Li et al., 2010).

As rates are calculated for all the infections and each organism, the ISPD guidelines in 2010, highlighted two ways of calculating peritonitis rates. Firstly, peritonitis rates should be calculated as the number of infections by organism for a time period, divided by dialysis-years' time at risk, and be expressed as episodes per year (Isla et al., 2014; Li et al., 2010). Secondly, the peritonitis rates should be calculated as the total number of patient-months on PD divided by the number of episodes of peritonitis and expressed as the number of episodes per year at risk and/or the number of months between episodes (Brown, 2012; Li et al., 2010). Li et al. (2010) reported that peritonitis rate could also be examined by calculating the percentage of patients who are peritonitis free per period of time and also by calculating the median rate for all the individual patient rates.

Peritonitis prevention or treatment

Previously, the guidelines of the ISPD focused on protocols to treat peritonitis rather than the prevention measures, however, the recent ISPD guidelines focuses attention on the prevention of peritonitis (Bender et al., 2006). Although, ISPD has recommended that in every PD program, effort should be made to prevent peritonitis, studies on preventing infections during PD are still limited both in quality and in quantity (Piraino et al., 2011). Boeschoten et al. (2006) indicated that an adequate catheter placement, proper exit side care, dedicated post operative catheter care and antibiotic prophylaxis are important measures for the prevention of peritonitis.

2.1.2 Possible risk factors

Age

The evidence regarding the effect of age on peritonitis is mixed. In several studies the age is identified as an important risk factor for an early onset of peritonitis (Ikabu et al., 2016a; Fan et al., 2014; Nishina et al., 2014). Okayama et al. (2011) found that aging (elderly people $>$ 80 years of age) is an essential risk factor of peritonitis in patients attending kidney dialysis center in Japan. Fan et al. (2014) also observed that elderly patients are at risk of having first episode of peritonitis. Ikabu et al. (2016a) found a significance association between age $>$ 40 years and the occurrence of peritonitis among South African CAPD patients. Moreover, Okayama et al. (2011) further emphasised that poor PD techniques because of advanced age may be the main reason for the findings and concluded that training programs and equipments aimed at elderly patients could help to reduce the associated risk.

Another study conducted in South Africa by Raaijmakers et al. (2010) found no correlation between age and peritonitis in children who are on PD. Interestingly, a study conducted from 1996 to 2005 in Canada by Nessim et al. (2009) reported that age is associated with higher peritonitis rate among patients who started PD between 1996 and 2000, but no significant association between age and peritonitis was found among patients who started PD between 2001 and 2005. They suggested that the variation in findings may have been caused by the differences in age cut-off values and limited statistical power (Nessim et al., 2009). They also speculated that the difference in findings may be caused by single-center studies and the era in which patients received PD, since many studies were conducted in different periods with different devices (Nessim et al., 2009). According to Li et al. (2007), nephrologist should no longer determine whether or not to offer the dialysis for elderly ESRD patients using only age. They indicated that renal team are faced with special challenges because of an increase of elderly dialysis patients.

Therefore, suggested that both social and medical issues have to be considered in providing elderly dialysis patients with appropriate care to avoid significant complications. Hence, from all these findings we can conclude that the influence of age as a factor for the outcome of peritonitis is debateable.

Gender

The literature concerning the role of gender in PD patients suggest that women are more likely to develop peritonitis than men (Kerschbaum et al., 2012; Lim et al., 2011; Kotsanas et al., 2007). Kotsanas et al. (2007) in Australia identified gender as the independent risk factor for peritonitis with women being almost twice as likely to have peritonitis compared to men. They indicated that one possible reason for their findings may be ascending infection from urogenital which in women are caused by an organism known as gram negative bacilli (Kotsanas et al., 2007). They concluded that the concurrent infection such as urinary or of the genital tract origin in female patients should be closely monitored (Kotsanas et al., 2007). Another study conducted by Kumar et al. (2014) in United State of America also found females to be significantly associated with peritonitis. They speculated that the relationship between females and peritonitis might be explained by BMI.

In contrast, Fan et al. (2014); Lobo et al. (2010) reported that men are associated with high risk of developing peritonitis as compared to women. Fan et al. (2014) showed that reasons for their findings may be due to more social interaction needs for men as compared to women which may cause the fluctuation of the living environment of the patient. They further pointed out that this may be due to the fact that the obedience of men was not good as compared to women. Lastly, women may have been more careful and had better personal hygiene as compared to men in their daily life.

Race

In the United States of America, Farias et al. (1994) evaluated the association between race and first peritonitis occurrence. They found that African-Americans had 50% higher risk of having peritonitis as compared to whites, even after the adjustment of some socioeconomic factors. Again, Korbet et al. (1993) found race as a significant risk factor for peritonitis in United States of America, particularly in black patients who had almost twice the peritonitis compared to white patients. They demonstrated that the observed increased peritonitis rate in black patients might be caused by racial differences in immunologic response (Korbet et al., 1993).

Similar findings emerged in South Africa by Ikabu et al. (2016a), who showed that black ethnicity stands a higher risk of developing peritonitis than other race groups. They demonstrated that a lower level of education, poor housing, lower number of qualified professionals, living in less developed environment, high cost of PD, lack of water and electricity supply might be responsible factors of high chance for peritonitis occurrence among black South Africans (Ikabu et al., 2016a). A study conducted by Martin et al. (2011), identified non-white race as a significant predictor of peritonitis in Brazilian PD patients. Indeed, they demonstrated that due to the adjusted confounding factors like the family's income and educational level, white race is most unlikely to have peritonitis. In contrast to the aforementioned findings, several studies found that race did not influence the peritonitis risk (Isla et al., 2014; Kumar et al., 2014; Kotsanas et al., 2007).

Dwelling

The issue concerning the role of rural and urban dwellers towards the development of peritonitis in PD patients is not well documented. Limited studies detailed the impact of rural and urban dwellers associated factors towards the development of peritonitis such as distance from hospital, individual income and education level

which are well-known risk factors of peritonitis. According to Isla et al. (2016) in South Africa, PD has the ability of being a preferred RRT option in developing countries because it might serve rural dwellers who sometimes live very far from the dialysis centers, often located in urban areas. They indicated that the effect of rural dwelling on infection related mortality among the dialysis patients in developing countries is known (Isla et al., 2016). Moreover, they showed that the option of PD as RRT among patients in rural dwelling is associated with infection related risks of death. They demonstrated that the reasons for their observations may be due to lack of finance in making these journeys from rural areas to urban areas where the centers are mostly located (Isla et al., 2016).

Xu et al. (2012) reported that PD patients who lived in undeveloped areas and regions of China were less educated and experienced a heavier burden of medical expenses. In addition, they found that low income earners, independently showed the highest risk of developing first peritonitis episode. They also indicated that patients in urban areas had comparable first peritonitis episode. Another study in China conducted by Wang et al. (2015) investigated the socioeconomic characteristics which could affect the outcomes of patients treated with PD. They found that most rural dwellers were younger with lower educational level, which might have led to lower awareness of chronic kidney disease.

With regards to Australian study by Lim et al. (2011) on the role of residential location on PD outcomes, they found patients residing in non-metropolitan location, particularly those in remote/rural areas to have higher risk of PD related complications including peritonitis. They observed that the inconsistency in mortality risk infection like peritonitis between rural and urban dwellers are due to reduced access to the specialized medical services. However, Chidambaram et al. (2011) in Canada discovered that residing in rural areas did not have an impact towards the risk of PD technique failure or death. They therefore, concluded that patients from rural areas should not be discouraged from starting PD as their findings indicated that patient's outcomes on PD are not altered by geographical location.

Catheter removal

The clinical practice guidelines for peritoneal access in Netherlands recommended that each PD center must have the ability to re-implant PD catheters and also urgently remove the PD catheters when necessary (Figueiredo et al., 2010). They further demonstrated that catheter removal can be performed either in the case of peritonitis or planned procedure, especially switching to HD or renal transplantation. The ISPD recommendations illustrated that every effort should be made to avoid peritonitis by replacing PD catheter for relapsing, recurrent, refractory exit site infections prior to the occurrence of peritonitis (Li et al., 2010). They indicated that in such cases, the PD catheter can be substituted as a single procedure. However, if peritonitis occurs, PD catheter should be removed and the patients be taken off from PD for a period of time (Li et al., 2010).

Dombros et al. (2005) indicated that every PD unit must analyse it's catheter survival and reasonable targets must include a catheter survival of at least 80% yearly with peritonitis rate of less than one episodes per 24 patient-months. They further showed that catheter removal for exit side infections could be considered if antibiotic treatment was unsuccessful (Dombros et al., 2005).

Piraino et al. (1991) investigated the patient's weight at the beginning of PD on the occurrence of peritonitis and catheter loss in Canada. They found that weight increase was associated with risk of catheter loss due to infections. Crabtree et al. (2003) examined the efficacy of silver-ion treated catheters in reducing PD related infections. They found no clinical effect between catheter with silver-ion and reduction of dialysis related infections. Several studies have also documented higher mortality rates when PD catheter removal is delayed. However, there is no evidence of studies which investigated the effect of catheter removal prior to initial or reoccurrence of peritonitis. Lonergan (2013) summarised the evidence available to assist with the timing of catheter removal impacting on the mortality rate and the return of patients to PD. Therefore, showing that the delay of catheter removal

decreases the return of patient to PD after peritonitis episode. Another study in the Northeast of Anatolia found that the placement of catheter using surgery increases a risk of developing peritonitis (Mustafa et al., 2015).

Education level

In many studies, educational level was found to be an independent and a strong predictor of peritonitis. However, the risks of technique failure due to educational level are different among many reports. According to a study conducted by Martin et al. (2011), education level is independently and negatively associated with first peritonitis episode in the Brazilian PD patients. They highlighted that these results were observed after the adjustment of some demographic and socio-economic factors as well as the relevant medical factors. These results confirmed the findings of Lobo et al. (2010), who discovered lower educational level as a risk factor for the development of peritonitis through multiple factors. In China, Fan et al. (2014) reported similar results which found lower educational level at the beginning of PD to be a major risk factor for the development of first peritonitis episode.

A recent report by Kim et al. (2017), observed that the lowest educational group has a higher chances of having peritonitis episode compared to a reference group. They speculated that these findings could be caused by differences in the effectiveness of training of various age groups (Kim et al., 2017). Furthermore, among the Taiwan's PD patients, Chern et al. (2013) found lower educational level as a major risk factor of peritonitis, irrespective of other factors such as hypoalbuminemia, age and gender. The impact of lower educational level may be associated with an decreased ability to learn the associated knowledge or operational procedures about PD (Fan et al., 2014). Additionally, Martín and Fernández (2006), showed that patients with learning problems had a limited time to the occurrence of first peritonitis episode.

Malnutrition: serum albumin

Serum albumin is a laboratory test which relates to protein energy malnutrition and its normal value of reference lies between 3.5 g/dl to 5 g/dl (Winterbottom, 2015). Serum albumin levels are more common among the patients with chronic kidney disease receiving PD due to the protein loss through the dialysate and inadequate dietary intake (Winterbottom, 2015). Malnutrition has consistently been found to be a potential risk factor of peritonitis and the evidence is clear from the literature.

In Scotland, Brown (2012) found serum albumin at the beginning of PD as an independent risk factor of peritonitis and it appears to be a significant predictor of peritonitis-free survival. Therefore, it was concluded that major risk factors such as lower RRF, age and gender could all potentially be associated with the low serum albumin (Brown, 2012). Fontan et al. (2005) observed peritonitis as a potential cause of mortality in patients who are on dialysis in Spain and identified malnutrition as an essential predictor of these complications. The findings were supported by Isla et al. (2014) in South Africa, who observed that poor nutrition could explain the high peritonitis rate and recommended the treatment of malnutrition to prevent the occurrence of peritonitis.

Santhakumaran et al. (2016) used hydration status as a potential risk factor for peritonitis in PD patients and found a positive relationship between over hydration and the high occurrence of peritonitis. However, the inclusion of nutritional status factors such as serum albumin in the multivariate analysis showed a reduced association between over hydration and peritonitis. They hinted that malnutrition could be the missing link between peritonitis and over hydration (Santhakumaran et al., 2016).

Boehm et al. (2005) investigated the potential risk factors for peritonitis in children undergoing PD as a primary RRT in Austria and could not find serum albumin as a potential predictor of peritonitis. This findings contradicted the findings of (Santhakumaran et al., 2016; Brown, 2012; Fontan et al., 2005). They emphasized

that this discovery may have been caused by a high percentage of children with nephrotic illness and concluded with similar remarks made by (Santhakumaran et al., 2016).

These remarks showed that malnutrition may indeed be an addictive major link between the residual renal output as an important risk factor for peritonitis and children. Abdu et al. (2011) noted that monitoring the nutrition status of patients who are on PD is important. They indicated that early identifications and good management of malnutrition could lead to a better nutritional status and patient's outcome. These were in accordance to the recommendations made by Lobo et al. (2010), which demonstrated that the measurement of serum albumin at the commencement of PD should be considered with reservations to infer the patients' nutritional statuses, which are proven to play a major role in the development of peritonitis.

Causative organisms

Research into the effects of gram negative and gram positive organisms in the development of peritonitis in patients who are on PD is limited. Studies which were conducted to assess the effect of these organisms reached the same conclusion. According to Van et al. (2014), peritonitis caused by gram-negative organisms are worse compared to the peritonitis caused by gram-positive organisms. A study conducted in India by Prasad et al. (2003) investigated the relationship between gram-positive and gram-negative organisms of 225 patients who are on CAPD in relation to the occurrence of peritonitis. They found peritonitis caused by gram-negative organisms to be higher compared to peritonitis caused by gram positive organisms, indicating that gram-negative organisms are highly associated with occurrence of peritonitis.

Moreover, Bunke et al. (1997) looked at the outcomes of a single organism (gram-negative against gram-positive) peritonitis in the United States of America. They

found that peritonitis caused by gram-negative organisms are worse compared to peritonitis caused by gram-positive organisms. Krishnan et al. (2002) discovered a high resolution rate of peritonitis caused by gram-positive organisms than peritonitis caused by gram-negative organisms. However, one of the studies conducted in Brazil by Lobo et al. (2010) assessed the incidence rate of peritonitis and its associated risk factors. They found that gram-positive and gram-negative organisms contribute similar proportions towards the development of peritonitis.

Body mass index

The association between peritonitis risk and obesity (BMI $>$ 25) has previously been reported. However, reports on the association of BMI and the survival in PD patients have been inconsistent in the literature. Several studies found BMI as a significant predictor of an early or first development of peritonitis episode and reoccurrence of peritonitis (McDonald et al., 2004). The studies conducted by Ikabu et al. (2016a); Isla et al. (2014) from South Africa revealed that obesity is associated with high incidence and early onset of peritonitis. Prasad et al. (2014), also found obese patients in India to have higher risks of developing peritonitis compared to patients with a normal BMI. Another study from Canada found that patients who are overweight at the beginning of PD have the same peritonitis rates as those who have normal weight (Piraino et al., 1991). They then concluded that, underweight patients at the beginning of PD are not at an increased risk of PD related infections (Piraino et al., 1991). According to Twardowski and Prowant (1996), obese patients are more at risk of the exit-site infection, which in itself is a risk factor of peritonitis.

However, the results from South Africa, India and Canada are contradicting those from Australia which revealed that patients who have the higher BMI survive more than twice as long as the patients whose BMI falls within a normal range (Johnson et al., 2000). They asserted that the chance of their findings could have occurred

by random probability of just 2% (Johnson et al., 2000). These findings contrast with the earlier studies which have found that being overweight has no effect or reduced survival on PD patients. Prasad et al. (2014) highlighted that the occurrence of peritonitis in patients with high BMI remains speculative. They attributed the speculation of BMI on the development of peritonitis to be loss of RRF in obese patients, poor wound healing and decreased resistance to infections in fat tissues (Prasad et al., 2014). According to McDonald et al. (2004), one possible reason for association between peritonitis and higher BMI could be colonisation and infection of the PD catheters after the patient's insertion which may result from an increased wounded area.

Haemoglobin level

Anemia is not a disease's name but a condition and it remains a major problem for chronic kidney disease patients who are on dialysis (Tsubakihara et al., 2010). It can be defined as a condition in which the Hb level is reduced beyond the normal reference levels (Winterbottom, 2015). According to the Guidelines on Renal Anemia in chronic kidney diseases by Tsubakihara et al. (2010) in the Japanese society for dialysis therapy, the targeted established guidelines for Hb level in PD patients should be at least 11 g/dL as the criterion for dose withdrawal or reduction. They have further demonstrated that Hb levels should be utilised as values of reference for the diagnosis of anemia and indicated that the normal reference range in healthy adults depends mainly on gender, age and race (Tsubakihara et al., 2010).

There have been very few studies conducted about the effect of lower Hb level towards the development of peritonitis. Abe et al. (2016) in Japan found that PD patients who have experienced peritonitis have significantly lower Hb levels. They demonstrated that this relationship between peritonitis and lower Hb level may have been a consequence of smaller administration periods of stimulating-

erythropoiesis agent in patients who had peritonitis than in those who did not (Abe et al., 2016). They further highlighted that these consequences could have contributed to the inadequate treatment of anemia in patients who had peritonitis (Abe et al., 2016). In South Africa, Isla et al. (2014) assessed the causes of peritonitis and concluded that Hb concentration is a major predictor of peritonitis. They have further demonstrated that the treatment of anemia could have a positive impact on PD outcomes by preventing the occurrence or reoccurrence of peritonitis (Isla et al., 2014).

Distance from hospital

Distance from the PD center is believed to be a contributory risk factor to the development of peritonitis in PD patients and consequently decreasing the patients survival on PD due to technique failure. Several studies identified significant association between distance from PD centers and peritonitis incidence rate (Sanabria et al., 2015; Figueiredo et al., 2014a; Xu et al., 2012). Cho et al. (2012) examined whether the distance from patient's residence and PD centers in Australia had an influence on the occurrence of peritonitis. They found that patients who are living more than 100 km far from the PD centers were associated with a higher risk of peritonitis. In addition, they reported that the distance to the closest PD center might influence patient outcome by compromised management. This compromised management includes tyranny distance, impacting upon access to medical care, delayed dialysate sample processing such as culture-negative organisms, and delayed diagnosis (Cho et al., 2012). An Australian study also discovered that the distance from the dialysis center is associated with an increased peritonitis rate (Gray et al., 2013).

Interestingly, a Brazilian study observed an association between the shorter distance to PD center and higher risk of first peritonitis episode (Martin et al., 2011). They attributed their findings to have been caused by greater availability of diagno-

sis resources in relation to PD center proximity (Martin et al., 2011). They further indicated that PD patients residing near to the PD center may in fact represent urban dwellers with poorer hygiene conditions (Martin et al., 2011). A Colombian study demonstrated that PD patients outcomes could be influenced by the distance of their home to the dialysis center (Sanabria et al., 2015).

However, in contrast to the aforementioned studies, Chidambaram et al. (2011) found that greater distance to PD center did not impact the risk of PD technique failure in Canada, which could have been caused by peritonitis. These findings are in line with (Isla et al. (2016, 2014); Raaijmakers et al. (2010)) in South Africa who discovered that distances from PD centers is not associated with peritonitis. Isla et al. (2014) demonstrated that the lack of association between long distance travelled to get to the dialysis unit and composite outcome is surprising due to high unemployment and lack of income among PD patients. These demonstration were in accord with Sanabria et al. (2015) which highlighted that traveling from remote or rural areas needed money and significant investment of time.

Water and electricity

There is no evidence of published studies in literature which directly evaluated the effect of water and electricity as risk factors of peritonitis. However, several studies conducted suggested that water and electricity supplies could be a major reasons why some factors such as housing status, dwelling, and race, among others are associated with peritonitis. A study conducted by Ikabu et al. (2016a) evaluated factors associated with peritonitis among CAPD patients and found black ethnicity to be associated with peritonitis. They emphasised that lack of electricity and water supplies among black South African might be a contributing factor of high probability for peritonitis (Ikabu et al., 2016a). Similarly, Jeda Chinchilla et al. (2016) investigated the risk factors for peritonitis in Guatemala and found housing conditions as a significant factor associated with peritonitis. They advised that, in

order to prevent peritonitis episode, homes should be provided with basic services such as water and electricity as well as satisfactory sanitation (Jeda Chinchilla et al., 2016). They further concluded that strategies in preventing peritonitis need to consider housing status, in relation to basic services.

Moreover, Peters (2014) believed that, although PD patients on home therapy can successfully manage to undergo PD dialysis at home. It is a good idea to start a home evaluation as early as possible. They demonstrated that potential problems such as suitable wiring for electricity devices and water supply should be evaluated (Peters, 2014). One study from South Africa investigated the peritonitis rate and all causes of composite-related peritonitis and found that having electricity and tap water at home is significantly different between those reaching composite outcome and those who did not (Isla et al., 2014). However, the study found that absence of tap water and electricity at home were not predictors of the outcome (Isla et al., 2014). Hence, as this aforesaid studies do not provide statistical evidence about the role of water and electricity towards the development of peritonitis. Further studies needs to be performed to offer statistical evidence about the role of water and electricity in relation to the development of peritonitis.

Employment

According to Isla et al. (2014), poverty continuous to be common in South Africa, particularly amongst the black African residents. They demonstrated a high unemployment rate (71.1%) with low level of income (50% earning below $180 per month) which might lead to poor access to tap water and poor access to health care services (Isla et al., 2014). This was in relation to Isla et al. (2016), who emphasised that PKDC patients are mostly unemployed and those accepted into the PKDC were offered a social grand of 1200 Rands per months from the provincial government. Research into the effect of unemployment status towards the development of peritonitis is very limited. A study conducted by Chow et al. (2005) in-

vestigated the effects of social factors on the development of peritonitis and a risk of hospitalisation. They discovered that employment status did not contribute to the development of peritonitis as well as affecting the risk of hospitalisation (Chow et al., 2005). Lobo et al. (2010) evaluated peritonitis incidence and outcomes in chronic PD patients. They found that family income was not associated with risk of peritonitis, which is itself a source of employment.

Seasons

According to Perencevich et al. (2008), recognition of the seasonal trend in medical institutions could improve infection diagnosis and prevention interventions. However, they indicated that there are few data available for seasonal variation towards the development of these infections (Perencevich et al., 2008).

Therefore, a number of studies have been conducted to examine the effect of seasons and the development of peritonitis. In Australia, Cho et al. (2011) studied the role of seasonal variation in PD associated peritonitis. They found the occurrence of first peritonitis episode in different seasons to be almost the same. They therefore, concluded that variation of seasons did not have appreciable impact on the overall clinical outcomes and PD peritonitis rates (Cho et al., 2011). Another study in United State of America by Perencevich et al. (2008), evaluated the summer peaks on the incidences of gram-negative organisms, which itself is an important predictor of peritonitis. They observed higher rates of gram-negative organism infections during summer months, due to elevated average monthly temperature (Perencevich et al., 2008).

Figueiredo et al. (2014b) examined the distribution of peritonitis in relation to seasonality and disease causing microorganisms in Brazil. They found gram-positive organisms to be responsible for the majority of the peritonitis occurrence with even distribution of the year. Therefore, with these findings, they concluded that seasonality had no effect on the occurrence of peritonitis. This conclusion was due to

failure to observed association of months of the year and peritonitis, although they have shown that peritonitis episodes in January and May was significantly higher in comparison to June.

A Brazilian study by dos Santos et al. (2013), investigated the compliance to hand hygiene (regarded as one of the precautions during the PD procedures) in relation to seasonality. They found that lower compliance in the first three months of the year (summer months) was significantly associated with peritonitis (dos Santos et al., 2013). They speculated that this may be due to reposition of less well trained staff and untrained employee who start working in February which could consequently lead to the development of peritonitis (dos Santos et al., 2013). Ikabu et al. (2016a) demonstrated the importance of climate and weather conditions by showing the higher variability of peritonitis incidence from different seasons over a year.

House type

Peters (2014) believed that PD patients on home therapy can successfully manage to undergo PD dialysis at home. However, indicated that starting home evaluation to manage potential problems like an unstable housing conditions as well as the evaluation of possible utility issue such as water supply as soon as possible could be a good idea (Peters, 2014).

The evidence about the effect of housing type is limited. Two studies defined houses as a brick if they were predominantly builded using bricks. Shacks have been defined as any housing that are build predominantly using any corrugated iron or discarded zinc (Isla et al., 2014; Katz et al., 2001).

With regards to South African literature on the housing type as an influencing factor in the likelihood of patient's peritonitis occurrence, Raaijmakers et al. (2010); Katz et al. (2001) are the authority in this area. Raaijmakers et al. (2010) looked at the

risk factors for the occurrence of peritonitis and found poor housing in PD patients to be significantly associated with development of peritonitis. They also observed a longer time-to-first peritonitis episode and significantly lower peritonitis rate in PD patients with good housing. However, Katz et al. (2001) evaluated the potential causes for peritonitis and found home conditions to have no influence towards the onset of peritonitis. Similar results were reported in relation to Raaijmakers et al. (2010) in Guatemala by (Jeda Chinchilla et al., 2016). They classified housing as fair (≤ 2 services) and good (3 services) and found that peritonitis rate was 2.5 times higher in PD patients with fair housing than in those with good housing (Jeda Chinchilla et al., 2016). They therefore concluded that housing is an appropriate risk factor related to the development of peritonitis (Jeda Chinchilla et al., 2016). They further indicated that in PD treatment, fair housing provided inappropriate conditions and emphasized that strategies in an attempt to prevent peritonitis episode should consider the housing status (Jeda Chinchilla et al., 2016). They advised that basic services in the home such as water services, adequate sanitation and electricity should be available to manage poor housing conditions (Jeda Chinchilla et al., 2016).

In sudan, Sayed et al. (2013) found no clear association between housing and the infection of peritonitis. They remark that the visits from PD nurses to patient homes enabled them to detect minor problems and find practical solutions (Sayed et al., 2013). Therefore emphasising that with proper instructions, most PD patients adapted their home environment for PD (Sayed et al., 2013). According to Bender et al. (2006), regular home visits could decrease the infection risk of peritonitis because it gives the opportunity for PD nurses to evaluate the environment and re-evaluate the patient's ability to undertake the dialysis procedures and to assess appropriate equipments the patient needs.

Diabetes

Winterbottom (2015) defined diabetes nephropathy as the presence of proteinuria of at least 0.5 $g/24hr$, hypertension, persistent albuminuria and a progressive decrease in kidney function leading to ESRD. There are two types of diabetes and Winterbottom (2015) demonstrated that although type one and type two diabetes mellitus are aetiologically distinct disorders, they share a common pathophysiological pathway through hypertension, hyperglycaemia as well as the consequent end-organ damage. Conflicting results have been reported about the differences between patients who are diabetic and those who are non-diabetic in terms of peritonitis incidence. For instance, the study conducted in Scotland by Brown (2012) found females with diabetes mellitus to have shorter peritonitis-free survival or a higher risk of having peritonitis. In an Australian study, McDonald et al. (2004) found an increase rate of peritonitis and worse peritonitis-free survival to be associated with diabetes mellitus.

In contrast to the aforementioned studies, several studies have not found any association between diabetes and risk of peritonitis (Rudnicki et al., 2010; Lobo et al., 2010; Kelecş et al., 2010; Kotsanas et al., 2007; Chow et al., 2005; Farias et al., 1994). Kerschbaum et al. (2012) emphasised that it is reasonable to regard diabetes mellitus as a risk factor for the development of peritonitis in patients who are on PD.

2.1.3 Related studies and their findings

Application of statistical techniques and results

The causes of peritonitis and the peritonitis rate of patients treated with CAPD in South Africa were assessed by (Isla et al., 2014). Univariate analysis was performed using independent t-test, Wilcoxon rank sum test and chi-square test to identify predictors of first peritonitis episode. It was observed that possible pre-

dictors of the outcome were not the prevalent adverse socio-economic or socio-demographic factors. Rather, factors such as Hb concentration, BMI and serum albumin were observed to be predictors of peritonitis. It was also found that the peritonitis rate decreased every year. Furthermore, from the findings of Martin et al. (2011) into the risk of first peritonitis episode in Brazil using Cox proportional hazard (PH) model, it was revealed that geographical factors, educational level, race and center size were associated with risk of first peritonitis.

The effect of RRF on the development of peritonitis was conducted by (Han et al., 2007). The study used Kaplan-Meier (KM) method, log rank test and multivariate Cox regression model to determine the risk factors and examine the difference of peritonitis-free period due to the presence of RRF and diabetes. It was discovered that RRF and diabetes were risk factors for first episode of peritonitis (Han et al., 2007). They also identified lower serum albumin level as a significant factor in the development of peritonitis after a loss or decline of RRF. These findings suggested that the preservation of RRF could be viewed as a protective strategy to reduce the occurrence of peritonitis. Logistic regression analysis was further applied in Scotland to investigate factors associated with increased risk of peritonitis. Several risk factors including lower RRF, older age, units, low serum albumin and diabetes mellitus in females were found to be significant for experiencing peritonitis. However, for a better PD practices and modifying of these risk factors, the study recommended that more studies should be conducted to come-up with possible reasons as to why some PD units are associated with higher risk of peritonitis (Brown, 2012).

Aging rather than diabetes mellitus, nutrition status or the efficiency of PD was shown to be an important risk factors of peritonitis in Japan using unpaired t-test, chi-square test and Cox PH model (Okayama et al., 2011). It was emphasised that poor PD technique due to advanced age could be the reasons for the findings. Cox PH model was used to determine factors associated with first episode of peritonitis in China (Fan et al., 2014). It was found that older age as well as other factors such

as lower educational level, male, and hypoalbuminemia were associated with the first episode of peritonitis.

The association between BMI and peritonitis rates among PD patients in New Zealand was investigated by McDonald et al. (2004) using Kruskal-Wallis test, chi-square test, analysis of variance and Wald test. It was found that higher BMI was associated with time-to-first peritonitis and more frequent occurrences, independent of other risk factors. It was also highlighted that the peritonitis-free survival was significantly less among patients with higher BMI and decrease monotonically with higher BMI. According to Twardowski and Prowant (1996), Obese patients showed to have more risk of exit-site infection, which itself is the most identified risk factor of peritonitis (Twardowski and Prowant, 1996; Piraino et al., 1991).

The outcomes of a single organism peritonitis in PD was investigated in Bunke et al. (1997) by comparing gram positives and gram negatives using chi-square test, fisher's exact test and t-test. They discovered that gram negative organism appears to be an important contributing factor to the development of peritonitis (Bunke et al., 1997). Furthermore, demonstrating that peritonitis caused by single non-pseudomonal gram negative organisms were proven to be a more serious infection than it was previously imagined (Bunke et al., 1997). According to Van et al. (2014), peritonitis caused by gram negatives organisms are worse outcomes compared to peritonitis caused by gram positive organism. Multiple logistic regression analysis was used by Chow et al. (2006) to assess risk factors of peritonitis and it was found that gram negative organisms were predictive independent risk factors for peritonitis.

2.2 Conclusion

This chapter started with a discussion on diagnostics, peritonitis rate and its prevention. The chapter also discussed some potential social, biological, socio-economic

and demographic factors of peritonitis. The studies into prognostic factors of peritonitis outcome have also been investigated in the literature with diverse results. The reasons for these diverse results could be linked to the following: first, studies were conducted in different countries, with different dialysis systems, hence the diversity in results. Second, difference in results may be linked to the difference in methodologies employed by these researchers. A number of studies conducted in literature have used linear regression, logistic regression and survival analysis techniques. However, those studies conducted in South Africa have made use of linear and logistic regression framework, which are both unsuitable for modelling time to first peritonitis because it ignores time. Therefore, this study applies survival analysis techniques to analyse time-to-first peritonitis of patients on PD. In conclusion, the literature review has backed the research hypothesis outlined in Chapter 1. Therefore, this research would like to go forward and statistically prove that there is a significant relationship between these variables with reference to PKDC, Pietersburg Provincial Hospital, South Africa.

Chapter 3

Methodology

3.1 Introduction

This chapter outlines the detailed description of statistical techniques used for this study and how the data was collected. The chapter starts by providing research design, source of data and a brief overview of non-parametric, semi-parametric and parametric survival analysis techniques.

3.2 Research design

This is a prospective study consisting of 159 patients treated with PD in Polokwane Kidney Dialysis Center (PKDC) at Pietersburg Provincial Hospital, Limpopo province, South Africa. All patients were followed on a monthly basis with a follow-up period between 2008 and 2015, inclusive. In this study, all patients had kidney disease and the outcome of interest was time-to-first peritonitis during the follow-

up period. Those patients who didn't show the event of interest, that is peritonitis, will be censored at the end of the study. Those who were lost to follow-up were also censored.

A number of categorical and continuous factors from social, demographic and biological perspective were recorded in the process. Some of the categorical factors that were recorded are present or absence of tap water at home, electricity, employment, dwelling, seasons, catheter removal, sex, race, level of education, house type, dialysate, and anuric at baseline, among others. Similarly, some continuous factors including BMI, cholesterol, distance from hospital, age at baseline, weight, serum albumin and Hb, among others were also recorded. SAS statistical package has been used for both data management and analysis purposes.

3.3 Source of data

The data set used in this study was prospectively collected from PKDC based on records of all kidney patients who were on PD between 2008 and 2015. The study's setting is Polokwane in the Limpopo province, South Africa. This center provides health services required by kidney patients to undergo dialysis. These services include catheter, dialysis solution and the training staff which teaches patients everything they need to know about PD, for instance, how to perform the exchange and cleaning of the catheter.

All kidney patients who initiated PD at the center were registered and relevant social, demographic and biological factors were recorded during the follow-up period. The study population consist of 159 patients' records and were evaluated on a monthly basis. The inclusion criteria is any patient who have been diagnosed with kidney failure and commenced PD between 2008 and 2015 at the center. At each follow-up period, clinical assessments, vital symptoms and biological tests were conducted by the staff team, and details of peritonitis event were also recorded.

To manage the data collection process, a list of patients on record every year was recorded into Microsoft Excel software with all aforementioned factors. However, patients who had started PD from other provinces and later joined PKDC after moving to Limpopo province where excluded in the study.

3.3.1 Dependent variable

The dependent variable is time-to-first peritonitis, and is coded 1 for all patients who experienced first peritonitis episode. Patients who did not experience first peritonitis episode were censored and coded 0. Nevertheless, these patients were also included in the analysis.

3.4 Survival analysis

Survival analysis model is defined as a collection of statistical procedures for the data analysis for which the outcome variable of interest is time until an event occurs. In this model, subjects are followed over a specified period of time until the event of interest occur (Gail et al., 2007). Survival analysis is simply time-to-event data, such as time-to-appearance of the disease, or time-to-death, among others. It was previously focused on predicting the mean time or probability of survival, and comparing the survival distributions of patients under different conditions. Survival analysis techniques are also used to analyse how long a patient lives or have been affected by a disease.

The three survival analysis techniques, namely non-parametric, semi-parametric and parametric are employed in this study. With these techniques, we can estimate the impact of factors on the risk of disease occurrence and also, examine the distribution of survival times (Fox, 2002). This risk of occurrence could change from year-to-year or may differ from one country to another with different factors

(Buis, 2006). Survival analysis model is defined by three main characteristics. Firstly, this model defines response variable as the waiting time until an event occurs. Secondly, patients are censored, in the sense that for some units, the event of interest has not occurred at the end of the follow-up period and finally, the independent variables whose effect on the waiting time we wish to access or control.

3.4.1 Time-to-event

The dependent (outcome) variable of interest is time until an event occurs. Therefore, the current study defines a dependent variable of interest as time-to-first peritonitis.

3.4.2 Censoring

The important feature of time-to-event data is censoring. Censoring occurs when there is an information about a particular patient's survival time, but we do not know the exact survival time. Hence, it is defined as the loss of observation on the lifetime variable of interest in the process of an investigation. Indeed, censored data contain a certain amount of information and therefore, it minimise the amount of information lost in survival analysis given the difficulty in collecting and obtaining information of an individual (Ohno-Machado, 2001). Some individual or subjects are censored simply because the event of interest does not take place before the study ends. Censoring is what distinguishes survival analysis from other fields of statistics, since it contains only a partial information about the random variable of interest (Miller Jr, 1983). Hence, is therefore, classified as left censoring and right censoring.

Right Censoring

Right censoring is defined as a case where all that is known, is that an individual is still disease free at the given time. This type of censoring occurs if an individual has entered the study but is lost to follow-up. The actual event time is therefore, placed somewhere to the right of the censored time along the time axis. It also occurs when the period of observation expires, or an individual is removed from the study before the event of interest occurs. It is often based on the assumption that an individual censored time is independent of the actual survival time (Liu, 2012). Right censoring is classified into three categories, namely, type I censoring, type II censoring and random censoring. However, the study will focus on random censoring.

Type I Censoring

This type of right censoring is usually related to predetermined observation period which is defined according to the research design. This simply means that the length of the censored time is equal to the length of the observation period. If we let $T_1, T_2, ..., T_n$ be a right censored survival variables by fixed constants $\alpha_1, \alpha_2, ..., \alpha_n$ and assume that $T_1, T_2, ..., T_n$ are independent of $\alpha_1, \alpha_2, ..., \alpha_n$. Then, if the observed sample consists of ordered pairs $(w_i, \delta_i), i = 1, 2..., n$ for each i, we have

$$w_i = min\{T_i, \alpha_i\}$$

$$\delta_i = \begin{cases} 1 & if \;\; T_i \leq \alpha_i \; (uncensored) \\ 0 & if \;\; T_i > \alpha_i \; (censored) \end{cases}$$

where α_i is a fixed censoring time and δ_i is the censoring indicator of T_i. In general, right censoring indicates a complete survival time interval, which is unknown (has been cut off), has been censored at the right side of the observed survival time interval. In simple terms, only some part of the observations would experience

a particular event of interest during the specified study interval and some would survive to the end of the follow-up period. Therefore, for those who survived until the end of the follow-up period, the only information known about them is that the actual survival time is located to the right of the end period.

Type II Censoring

This type of censoring is often encountered in the industrial applications. From n ordered failure times, only the initial $r(r \leq n)$ times are observed and others are censored. That is, subjects are included in the study at the same time and therefore, being followed-up until the predefined number of events is observed.

Type III Censoring (Random censoring)

Right censoring also occurs randomly at any time during the study. This type of censoring is called random censoring. Random censoring is defined as a case where patients enters the study at random times. Hence, if the sample of censoring times C_i $(C_1, C_2, ..., C_n)$, which are assumed to be independent and identically distributed ($i.i.d$), each with a certain distribution, where each C_i is associated with T_i. We can observe that if $(w_1, \delta_1), ..., (w_n, \delta_n)$, where for $i = 1, ..., n$, then

$$w_i = min\{T_i, C_i\}$$

$$\delta_i = I(T_i \leq C_i) = \begin{cases} 1 & if \;\; T_i \leq c_i \; (uncensored) \\ 0 & if \;\; T_i > c_i \; (censored) \end{cases}$$

where δ_i, for $i = 1, ..., n$ contains the censoring information and $w_i, ..., w_n$ are $i.i.d$ with some distribution. Random censoring arises in medical application with clinical trials (Miller Jr, 2011). Therefore, this type of censoring, mainly occurs in clinical trials were patients may enter the study at different times. It differs with type I censoring due to the fact that its censored times are not fixed, since patients enter the

study at different times. Hence, the censored survival time for random censoring is measured as the time distance from the time of entry into the study to the time when random censoring occurs.

Left Censoring

Left censoring is a type of censoring where an individual or subject's true survival time is less than or equal to an individual or subject's observed survival time (Gail et al., 2007).

3.4.3 Survivor function

Suppose that T is a random variable denoting the survival time. Therefore, T is defined as a waiting time until an event occurs and is also regarded as a non-negative continuous random variable with the probability density function (pdf) $f(t)$ and cumulative distribution function (cdf) $F(t)$, given by

$$F(t) = P(T \leq t) \tag{3.1}$$

which provides the probability that an event has occurred at time t. Therefore, the survival function ($S(t)$) is defined as the probability that the subject or an individual survives longer than the specified time t, and it is a non-increasing function. That is, it takes value zero at infinity and value one at zero and it is given by

$$S(t) = P(T > t) = 1 - F(t) = \int_t^\infty f(x)dx \tag{3.2}$$

which is the complement of the cdf. Hence, $S(t)$ gives the probability that the random variable T exceeds the specified time t. $S(t)$ is very useful for comparing the survival progress of at least two groups.

3.4.4 Hazard function

Consider a random variable T, which represents the time until an event occurs and let the hazard function of T, which is defined as an instantaneous potential per unit time for an event to occur, given that the event has survived up to time t, be denoted by $h(t)$. Therefore, the hazard function ($h(t)$) is given by

$$h(t) = \lim_{\triangle t \to 0} \frac{P(t \leq T < t + \triangle t | T \geq t)}{\triangle t} = \frac{f(t)}{1 - F(t)} \tag{3.3}$$

where the limit at small time interval approaches zero with the numerator expressing the conditional probability, which gives the probability that an individual or subject's survival time will lie in the interval between t and $t+ \triangle t$. Hence, this equation may also be expressed as

$$h(t) = \frac{f(t)}{S(t)} \tag{3.4}$$

In other words, based on equation 3.4, $h(t)$ is a rate rather than the probability or a density, where $S(t), F(t), f(t)$ are the survivor function, cdf and the pdf of T, respectively (Jain and Vilcassim, 1991). In this sense, the hazard is a measure of risk. Generally, from equation 3.4, there is a relationship between the three specifications $S(t), h(t), f(t)$ of the probability distribution of T and, its also clear that $h(t)$ is finite (no upper bound) and nonnegative. Therefore, the relationship between $h(t)$ and $S(t)$ can be express equally as either of the following equations,

$$S(t) = e^{[-\int_0^t h(u)du]} \tag{3.5}$$

which describes how $S(t)$ can be written in terms of an integral involving $h(t)$ and

$$h(t) = -[\frac{dS(t)/dt}{S(t)}] \tag{3.6}$$

which describes how $h(t)$ can be written in terms of derivative involving $S(t)$. In short, $h(t)$ gives a useful description of the risk of failure at any time point.

3.4.5 Non-parametric estimation of survivor function

In survival analysis, it is at all times worthy to present graphical or numerical summaries of survival times of individuals. Comparison of survival curves will help to determine whether different samples (groups) could have arisen from the same survivor functions. Therefore, Kaplan Meier estimator and log rank test will be consider to evaluate the equality of survivor functions at some specified time t. These approaches are said to be non-parametric methods, since they need no assumptions about the distribution of T.

Kaplan Meier estimator or Product Limit formula

The Kaplan Meier estimate of the $S(t)$ is a non-parametric method of estimating $S(t)$, based on $i.i.d$ survival times that can be non-informatively right censored data. The method is widely used in many health science studies, and it has been identified as an important tool in the analysis of censored time-to-event data (Miller Jr, 1983) due to its capability in estimating the survival distribution for censored data. It is one of the best options to be used in measuring the fraction of subjects living for a certain amount of time after treatment (Goel et al., 2010). Kaplan Meier estimator has been one of the key statistical methods for analyzing censored survival data and it is defined as follows.

Suppose that $T_1, T_2, ..., T_n$ are $i.i.d$ survival times with $S(t)$ and the censoring times $C_1, C_2, ..., C_n$, which are independent of T_i. Now, if the observations are denoted by (X_i, δ_i) for $i = 1, 2, ..., n$, where $X_i = T_i \wedge C_i$ and $\delta_i = 1_{(T_i \leq C_i)}$.

Then, suppose that $F(t)$ is discrete with mass point at $0 \leq t_1 < t_2 < ...$, and therefore, define the discrete $h(t)'s$ as $h_1 = P(T = t_1)$ and $h_j = P(T = t_j \mid T >$

$t_{j-1})$ for $j > 1$. Then, if $t\epsilon(t_j, t_{j+1})$. By definition we have

$$\begin{aligned} S(t) = P(T > t) &= P(T > t_j) \\ &= P(T > t_j \mid T > t_{j-1}) \times P(T > t_{j-1}) \\ &= P(T > t_j \mid T > t_{j-1}) \times P(T > t_{j-1} \mid T > t_{j-2}) \times P(T > t_{j-1}) \\ &= (1 - h_j) \times (1 - h_{j-1})...(1 - h_1) \\ &= \prod_{i=1}^{j} (1 - h_i). \end{aligned}$$

Similarly, if we let $f_1 = h_1$, with $j > 1$, by definition we have that

$$f_j = h_j = P(T = t_j) = h_j \times \prod_{i=1}^{j-1} (1 - h_i).$$

Now, by making an inference about F based on likelihood function $L(F)$ corresponding to (X_i, δ_i) for $i = 1, 2, ..., n$. Then, we have

$$L(F) = \prod_{X_i:\delta_i=1} f(X_i) \times \prod_{X_i:\delta_i=0} [1 - F(X_i)]$$

and by substituting h_j we get

$$L(F) = \prod_{j=1} h_j^{d_j} \times (1 - h_j)^{n_j - d_j} \tag{3.7}$$

for

- $0 \leq h_j \leq 1$.
- $n_j = \sum_{i=1}^{n} 1_{(X_i \geq t_j)}$ = number of individual at risk (alive) just before time t_j, including those who will die at time t_j.
- $d_j = \sum_{i=1}^{n} \delta_i \times 1_{(X_i \geq t_j)}$ = number of failure (peritonitis) at time t_j.

Hence, the maximizing solution is seen to be for $n_j > 0$, $\widehat{h_j} = \frac{d_j}{n_j}$ and therefore

resulting in,

$$\widehat{S}(t) = \begin{cases} 1 & \text{if} \quad t < t_1 \\ \prod_{i=1}^{j-1}(1-\widehat{h_i}) \text{ if } t_j \leq t < t_{j+1}. \end{cases} \tag{3.8}$$

Moreover, Pollock et al. (1989) demonstrated that the specification of a particular parametric continuous distribution is not required.

Goel et al. (2010) indicated three assumptions considered when computing Kaplan Meier analysis and are as follows.

- Firstly, the survival probabilities are assumed to be the same for all individuals recruited early and late in the study.
- Secondly, it is assumed that at any time, individuals or subjects who are censored have the same survival prospects as those who are still in the study.
- Thirdly, it is assumed that the event occur at the specified time.

This estimate can also be expressed as a product limit formula which is given by

$$\widehat{S}(_{t_{j-1}}) = \prod_{i=1}^{j-1} \widehat{P}(T > t_{(i)} | T \geq t_{(i)}) \tag{3.9}$$

Equation 3.9 is good in describing the survival of group of individuals, and it also produces the survival estimates of different strata and require arbitrary discretization of continuous variables (Ohno-Machado, 2001).

Greenwood formula

Consider the likelihood formula in equation 3.7, then it follows that a large sample variance of $\widehat{h_i}$ is given by the usual binomial formula

$$var(\widehat{h_i}) = \frac{\widehat{h_i} \times (1-\widehat{h_i})}{n_i}$$

where $\widehat{h_i} = \frac{d_i}{n_i}$, with d_i following a binomial distribution, which is approximated by a normal distribution and $\widehat{h_i}$, $i = 1, 2, ...$ are independent in large samples. Now, taking the log of $\widehat{S}(t)$, we get

$$\log[\widehat{S}(t)] = \log[\prod_{i=1}^{j}(1 - \widehat{h_i})] = \sum_{i=1}^{j} \log(1 - \widehat{h_i})$$

Again, by taking the variance of $\log[\widehat{S}(t)]$, we obtain

$$var(\log[\widehat{S}(t)) = var(\sum_{i=1}^{j} \log(1 - \widehat{h_i})) = \sum_{i=1}^{j} var[\log(1 - \widehat{h_i})]$$

Hence, note that since d_i approximate binomial distribution, which is approximated by a normal distribution, then $\widehat{h_i}$ will also approximate normal distribution and allow us to apply the Delta (δ) method. Hence, by applying (δ) method we obtain

$$\begin{aligned} var(\log[\widehat{S}(t)]) &= \sum_{i=1}^{j} var(\widehat{h_i}) \times (\frac{1}{1 - \widehat{h_i}})^2 \\ &= \sum_{i=1}^{j} \frac{\widehat{h_i} \times (1 - \widehat{h_i})}{n_i} \times \frac{1}{(1 - \widehat{h_i})^2} \\ &= \sum_{i=1}^{j} \frac{\widehat{h_i}}{n_i} \times \frac{1}{(1 - \widehat{h_i})} \\ &= \sum_{i=1}^{j} \frac{\widehat{h_i}}{n_i \times (1 - \widehat{h_i})} \end{aligned}$$

Therefore, substituting $\widehat{h_i} = \frac{d_i}{n_i}$, we get

$$var(\log[\widehat{S}(t)]) = \sum_{i=1}^{j} \frac{d_i}{n_i \times (n_i - d_i)}$$

Now, since $\widehat{S}(t) = e^{\log[\widehat{S}(t)]}$ and if we have to use δ method again, this time to get the variance of the survivor function from the variance of its log, we then get a

Greenwood formula which is given by

$$\begin{aligned} var(\widehat{S}(t)) &\approx var(\log[\widehat{S}(t)]) \times (e^{\log[\widehat{S}(t)]})^2 \\ &= \widehat{S}(t)^2 \times var(\log[\widehat{S}(t)]) \\ &= \widehat{S}(t)^2 \times \sum_{i=1}^{j} \frac{d_i}{n_i \times (n_i - d_i)} \end{aligned}$$

and its use is to get the approximate confidence interval.

Log Rank Test

Log rank test is the most powerful rank invariant test when the censoring is applied equally to both groups (samples) (Crowley, 1974). Peto and Pike (1977) identified the test as a local power for detecting differences in failure rates. It is the most large sample chi-square test used for the comparison of the Kaplan Meier curves (Gail et al., 2007). It is used to test whether the difference between survival times among two groups is statistically different or not, but fails to allow to test the effect of some independent variables.

Now, suppose that $t_1 < t_2 < ..... < t_n$ are distinct times of observed failures and define the number of events and the number of movements at t_i in the following 2×2 contingency table.

Table 3.1: 2×2 contingency table

Treatment group	Events d_i	No events	At risk at t_i
K_1	d_{1i}	$n_{1i} - d_{1i}$	n_{1i}
K_2	d_{2i}	$n_{2i} - d_{2i}$	n_{2i}
Total	d_i	$n_i - d_i$	n_i

where,

- n_{ki} = number of persons in group k who are at risk at $t_i (k = 1, 2), i = 1, 2, ..., n$.

- $n_i = n_{1i} + n_{2i}$ =number at risk at t_i (both groups).
- d_{ki} =number in group k who fail (uncensored) at t_i.
- $d_i = d_{1i} + d_{2i}$ = total number of failures at t_i.

So, if the hypothesis of no association holds, that is:

$H_0 : F_1(t) = F_2(t)$ or $H_0 : S_1(t) = S_2(t)$,

then, the four marginal totals should all be fixed, say a single element d_{2i} and with conditioning and assuming H_0, d_{2i} has a hypergeometric distribution. Meaning that the hypergeometric probability of having d_{2i} in n_{2i} given fixed values of n_i, n_{2i} and d_i can be written as

$$P(d_{2i}) = \frac{\begin{pmatrix} d_i \\ d_{2i} \end{pmatrix} \begin{pmatrix} n_i - d_i \\ n_{2i} - d_{2i} \end{pmatrix}}{\begin{pmatrix} n_i \\ n_{2i} \end{pmatrix}},$$

for $d_{2i} = max(0, d_i - n_{1i}), ..., min(d_i, d_{2i})$. Hence, the mean and variance of d_{2i} under H_0 are

$$E_i = \frac{d_i n_{2i}}{n_i}$$

and $var(d_{2i}, n_i, n_{2i}, d_i) = \frac{n_i - n_{2i}}{n_i - 1} \times \frac{d_i n_{2i}}{n_i} \times (1 - \frac{d_i}{n_i}) = \frac{n_{1i} \times n_{2i} \times d_i (n_i - d_i)}{n_i^2 (n_i - 1)}$. Now, if we define $O_i = d_{2i}$, then fisher's test would tell us to consider extreme values of d_{2i} as evidence against H_0, thus defining $O = \sum_{i=1}^{n} O_i =$ number of failures in group 2 , $V = var = \sum_{i=1}^{n} V_i, E = \sum_{i=1}^{n} E_i$ and let $\frac{O-E}{\sqrt{V}} \sim N(0, 1)$ as $O - E$ tends to be normally distributed with increasing sample and its standardised form has mean zero and variance one. Therefore, the $Z - score$ can be derived for testing independence of

survival groups with test statistics defined by

$$Z = \frac{\sum_{i=1}(O_i - E_i)}{\sqrt{\sum_{i=1} V_i}}.$$

Therefore, under H_0, it is argued that $Z \sim N(0,1)$ or $Z^2 \sim X_1^2$ and if $Z = Z_\alpha, H_0$ should be rejected. Also, if $H_0 : S_1(t) = S_2(t)$, it should be accepted with the conclusion that survival and group are independent. Moreover, if the total number of observed events are large, the standard normal is converted to X^2 distribution, which is given by

$$X^2_{logrank} = \frac{\sum_{i}^{n}(O_i - E_i)^2}{\sum var(d_{2i})} \tag{3.10}$$

where $X^2_{logrank}$ is the log-rank statistics and, $O_2 - E_2$ is the sum of observed minus the expected score for one of the groups being compared with one degrees of freedom. $X^2_{logrank}$ will be used to determine the overall difference between the two survival curves being compared. $X^2_{logrank}$ is most powerful, if the odds ratios (or hazard ratios) among the samples are constant (called "proportional hazards") over time and the departure from the proportional hazards can be checked by examining the estimated survival curves.

3.4.6 Cox Proportional Hazard Model

The Cox Proportional Hazard (PH) model has become very popular for the analysis of failure time observations (Lin et al., 1993). The model includes the product of two quantities at hazard time t, which are the baseline hazard function $h_0(t)$ that is independent of covariates $X^{'}s$ and the exponential for the linear sum of $\beta_i X_i$, denoted by $e^{\sum_{i=1}^{p} \beta_i X_i}$, which is independent of time t. The PH model enables the testing of the effect of independent variables on survival times of different groups

of individuals or subjects and, is given by

$$h(t|X) = h_0(t) \times e^{\sum_{i=1}^{p} \beta_i X_i} \tag{3.11}$$

where $h_0(t)$ represents $h(t)$ for an individuals with factor values all equal to zero, $X_i = (x_1, x_2, ..., x_p)$, which denotes the collection of factors for the i^{th} individual and the estimates, $\beta_i = (\beta_1, \beta_2, ..., \beta_p)$, which denotes a collection or vector of p unknown coefficients.

The absence of a parametric form of a survival distribution gives Cox model its other name, known as semi-parametric model. This is because the only parameters to estimate in the model are those describing how the predictors affects the hazard. The PH model reduces to

$$h(t, x) = h_0(t) \times e^{\beta_i^{|} x} \tag{3.12}$$

for a single binary factor. Then, if the variable is binary, $h(t)$ is $h(t, x = 0) = h_0(t)$ when the variable has the value of 0 and $h(t, x = 1) = h_1(t) = h_0(t) \times e^{\beta}$ when the variable has the value of 1.

The Cox PH model explore the relationship between the survival of patients and several covariates (Kartal, 2015). Hence, it is used to estimate the effects of parameters. This model closely approximate the correct model even though $h_0(t)$ is not specified. Therefore, the good estimates of hazard ratios of interest, regression coefficients and adjusted survival curves can be obtained. Furthermore, the measure of effect (hazard ratio) is computed without having to estimate the baseline intensity. Even thought the $h_0(t)$ is not specified, survivor and hazard curves can be estimated.

The model assumes proportional hazards between the values of predictors, irrespective of how fundamental the hazard could change over time. Since Cox PH model is constructed completely around this assumption, if it happens that it is in-

valid for a set of predictors in the data set, the Cox PH model must not be used on that data set. That is, if the hazards cross for different strata, the PH assumption is questionable. But if they are parallel, the PH assumptions is satisfied. However, there are options available if the assumption is not valid, that is to analyse by stratifying on the exposure variable, without fitting any model and instead, KM curves for each exposure group can be obtained separately.

Hazard Ratio

Hazard ratio (HR) depends on the follow-up time (Royston and Parmar, 2013) and it is defined as the ratio of the predicted $h(t)$ under two different values of a predictor variable. HR is denoted by

$$\widehat{HR} = \frac{\widehat{h}(t, X^*)}{\widehat{h}(t, X)} \tag{3.13}$$

where $X^* = (X_1^*, X_2^*, ..., X_p^*)$ indicates the set of predictors of an individual and $X = (X_1, X_2, ..., X_p)$ indicates the set of predictors for another individual. Furthermore, HR can aslo be expressed as

$$\begin{aligned} \widehat{HR} &= \frac{\widehat{h_0}(t) e^{\sum \widehat{\beta}_i X_i^*}}{\widehat{h_0}(t) e^{\sum \widehat{\beta}_i X_i}} \\ &= e^{\sum_{i=1}^{p} \widehat{\beta}_i (X_i^* - X_i)} \end{aligned} \tag{3.14}$$

Therefore, HR is independent of time t (constant over time t) and the hazard rates are proportional, hence a PH model. HR compares the hazard of having an event with covariate value X^* to the hazard of having an event with covariate value X. A HR less than one means that the event is less likely to occur, and a HR greater than one means the event is more likely to occur.

Cox's Partial Likelihood

The focus of Cox model is to estimate the regression coefficients using partial likelihood (Klein et al., 2013). That is, to efficiently estimate the parameters β_i when the functional form of $h_0(t)$ is not known. Therefore, since

$$\frac{hazard\ rate\ at\ t_i\ for\ individual\ i\ with\ covariates\ x_i}{\sum_{j \epsilon R(t_i)} hazard\ rate\ at\ t_i\ for\ individual\ j}$$

as both the numerator and denominator are in terms of $h(t)$ given that $\Delta \longrightarrow 0$, where the individuals which are at risk at an instant before time t_i are denoted by $R(t_i)$. Therefore, the probability that a specified individual has an event at time t_i, given that one of the individuals in $R(t_i)$ has an event at time t_i is given by

$$\begin{aligned}\frac{h(t|x_i)}{\sum_{j \epsilon R(t_i)} h(t|x_j)} &= \frac{h_0(t_i)e^{x_i^{'}\beta}}{\sum_{j \epsilon R(t_i)} h_0(t_i)e^{x_j^{'}\beta}} \\ &= \frac{e^{x_i^{'}\beta}}{\sum_{j \epsilon R(t_i)} e^{x_j^{'}\beta}}\end{aligned}$$

As a result, we get the joint likelihood for β, which is specified as a product over all t_i values, and it is indicated in equation 3.15 below.

$$L(\beta) = \prod_{i=1}^{n} [\frac{e^{x_i^{'}\beta}}{\sum_{j \epsilon R(t_i)} e^{x_j^{'}\beta}}]^{\delta_i} \tag{3.15}$$

Equation 3.15 shows the only parameter to be estimated by Cox model, with $L(\beta)$ representing an incomplete likelihood function and δ_i denoting a censoring indicator such that if t_i is an event, $\delta_i = 1$ and if t_i is a censored time, $\delta_i = 0$ (Liu, 2012).

Furthermore, if the observation is censored, that is when $\delta_i = 0$, then $L(\beta) = 1$, this means that the product of the conditional probability for all right censored

individuals are 1, indicating that there is no any contribution to the $L(\beta)$ for the right censored cases, because it shows only probabilities for failed subjects. Hence, they can be unaccounted for in $L(\beta)$ without influencing the total $L(\beta)$. $L(\beta)$ is the conditional probability of an event occurrence given the risk set $R(t_i)$ if $\delta_i = 1$. The above incomplete likelihood function is called partial likelihood function in survival analysis since there is no information about $h_0(t)$ which contributes to it.

Cox proposed the following partial likelihood for the parameter β, which simplify equation 3.15 by only multiplying the conditional probabilities over all events and get

$$L(\beta) = \prod_{i=1}^{d} \frac{e^{x_i^{'}\beta}}{\sum\limits_{j \epsilon R(t_i)} e^{x_j^{'}\beta}} \tag{3.16}$$

where d is the distinct failure times (total number of events) t_i for $i = 1, ..., d$, which are ordered event time and i denotes the subscript of subjects who had an event at t_i. Then, equation 3.16 still denotes the conditional probability that a patient experiences an event at time t_i given all the patients exposed to that risk at t_i.

This general justification of $L(\beta)$ as part of the full likelihood was provided by Cox in 1975, as a part that happens to contain most of the information about β. This justification is valid even with time varying covariates (Rodrıguez, 2005).

3.4.7 Estimation of Cox hazard model with tied of survival times

The Cox PH model assumes that hazards should be continuous and that ties should not be heavy. When times in continuous time model are grouped, ties in failure times can be observed. However, in $L(\beta)$ ties are not allowed, because $L(\beta)$ is valid only for data which are not grouped (Bêlavŝková et al., 2013). But when ties do happen, there are some proposed modifications to partial likelihood for the adjustment of ties and they are as follows.

The Breslow method

Suppose that the events occurs at D distinct times $t_1, t_2, ..., t_D$. Then, the Breslow partial likelihood function is given as

$$BreslowL(\beta) = \prod_{i=1}^{D} \frac{e^{\beta' s_i}}{(\sum_{j \epsilon R(t_i)} e^{\beta' x_j})^{d_i}} = \prod_{i=1}^{D} \frac{e^{\beta' \sum_{j \epsilon D_i} x_j}}{(\sum_{j \epsilon R(t_i)} e^{\beta' x_j})^{d_i}} \tag{3.17}$$

where D_i is denoted as a set of all subjects who failed at time t_i, d_i is the total number of failures at time t_i and $s_i = \sum_{j \epsilon D_i} x_j$ is the sum of covariates values all over all subjects in the set D_i (Xin, 2011).

Breslow's method provides a much simpler approach to estimate the survival function in Cox model. Clearly, the method accounts for the contribution of d_i events simply by multiplying the conditional probabilities over all events at t_i. Therefore, if at each survival times the d_i is small or the number of patients at risk n_i is large, meaning that the ratio of $\frac{d_i}{n_i}$ is small, then Breslow method's approximation should work well, meaning that the approximated $L(\beta)$ should be very close to the exact $L(\beta)$.

However, the method can perform poorly if the data are heavily tied (happen mostly when the data are either interval censored by study design or the time scale is truly discrete). The method can also cause a severe bias for the datasets with large fraction $\frac{d_i}{n_i}$ and the approximation can be poor if the conditions are not satisfied. Another issue with Breslow method is that it considers each of the events at a given time as distinct and allows all failed subjects to contribute fully to the risk set. Hence, Efron's method was suggested.

The Efron's method

According to Xin (2011), the Efron's method also deals with ties by approximating the $L(\beta)$ and is given by

$$\begin{aligned} EfronL(\beta) &= \prod_{i=1}^{D} \frac{e^{\beta' s_i}}{\prod_{j=1}^{d_i} (\sum_{k \epsilon R(i)} e^{\beta' x_k} - \frac{j-1}{d_i} \sum_{k \epsilon D(i)} e^{\beta' x_k})} \\ &= \prod_{i=1}^{D} \frac{e^{\beta' \sum_{j \epsilon D_i} x_j}}{\prod_{j=1}^{d_i} (\sum_{k \epsilon R(i)} e^{\beta' x_k} - \frac{j-1}{d_i} \sum_{k \epsilon D(i)} e^{\beta' x_k})} \end{aligned} \tag{3.18}$$

This method allows a partial contributions to the risk set for each of the members that fail at t_i. Furthermore, it reduces the weight of the denominator by introducing ordering into the $L(\beta)$ and is regarded as a good approximation for the discrete hazard model. Hence, obtaining the parameter estimates on the basis of the above defined approximation is not particularly complicated.

However, Borucka (2014) pointed out that the estimators resulting from $BreslowL(\beta)$ and $EfronL(\beta)$ might be biased. This method is preferable when the sample size is small either from the outset or due to heavy censoring.

3.4.8 PH Model Assumption Checking

PH model assumes the nonlinear relationship between the predictor variables and the hazard function. Therefore, since the nonlinear relationship is assumed, the hazard ratio comparing any two subjects or individuals is constant over time when the predictor variables are not different over time. This assumption is called the PH assumption. If the data fails to follow this assumption, the results of the study using PH model would end up being invalidated.

Graphical Approach

If we plot the estimated log-log survival curves for individuals on the same plot using Cox PH model, the two curves should be approximately parallel. That is, by plotting the estimated log-log (survival) against the survival time for the two groups of individuals we would observe parallel curves if the PH assumption holds.

Now, suppose that the Cox PH of a $S(t)$ can be obtained by a relationship of the $S(t)$ and $h(t)$, where

$$S(t, X) = [S_0(t)]^{e^{\sum_{i=1}^{p} \beta_i x_i}} \tag{3.19}$$

with $X = (x_1, ..., x_p)'$ indicating values of independent variables for a particular subject. Hence, when the logarithm is applied, we easily get

$$ln[-lnS(t, X)] = \sum_{i=1}^{p} \beta_i x_i + ln[-lnS_0(t)] \tag{3.20}$$

Furthermore, the difference in log-log curves corresponding to two different subjects having factors $X_1 = (x_{11}, ..., x_{1p})$ and $X_2 = (x_{21}, ..., x_{2p})$ is denoted by

$$ln[-lnS(t, X_1)] - ln[-lnS(t, X_2)] = \sum_{i=1}^{p} \beta_i (x_{1i} - x_{2i}) \tag{3.21}$$

which does not involve time t or does not depend on t.

The parallelism of the plotted log-log survival plot provides us with a graphical way for checking the Cox PH assumption. However, the approach performs poorly when assessing PH assumption for continuous factors and categorical factors with many subcategories or levels (Jiezhi, 2009).

Moreover, assessing PH assumption based on log(-log (survival)) against the survival time as well as KM-curves is not enough to be sure about proportionality, since this approaches are univariate analysis, and fail to indicate whether the hazards could still be proportional when Cox PH model incorporates many factors at

the same time. Therefore, some approaches that could be used to validate assumption of proportionality if factors are included together in the Cox PH model were proposed. One such approach is the supremum test for proportional hazard assumption.

Supremum Test for proportional hazards assumption

The standardized test statistics

$$\widetilde{U}_m^*(t) = sup[(I^{-1}(\widehat{\beta})_{mm})^{\frac{1}{2}}\widetilde{U}_m^*(\widehat{\beta}, t)], m = 1, ..., M.$$

with $I^{-1}(\widehat{\beta})_{mm}$ representing the diagonal elements in the inverse of the respective observed information matrix. This $\widetilde{U}_m^*(t)$, under the null hypothesis that the proportionality assumption hold, it is referred as a special case of Weiner stochastic process given by

$$W(t, z) = \sum_{i=1}^{n} f(Z_i)I(Z_i \leq z)\widehat{P}_i(t)$$

with $z = \infty$ and $f(.) = ..$ Therefore, given the Taylor series expansion, $\widetilde{U}_p^*(t)$ is approximated by

$$\begin{aligned}\widetilde{U}_p^*(t) = (I^{-1}(\widehat{\beta})_{pp})^{\frac{1}{2}}[&\sum_{l=1}^{n} I(T_l \leq t)\delta_l(Z_{pl} - \overline{Z_p}(\widehat{\beta}, t))G_l \\ &- \sum_{k=1}^{n}\int_0^t Y_k(u)e^{z_k^{'}\widehat{\beta}}Z_{pk}(Z_k - \overline{Z}(\widehat{\beta}u))^{'}d\widehat{\Lambda_0} \\ &\times I^{-1}(\widehat{\beta})\sum_{i=1}^{n}\delta_l(Z_l - \overline{Z}(\widehat{\beta}, T_l)G_l)]\end{aligned} \tag{3.22}$$

with $\overline{Z_p}(\widehat{\beta}, t)$ representing the p^{th} component of $\overline{Z}(\widehat{\beta}, t)$, G_l is a standard normal variables that are independent of the triple (T_l, Z_l, δ_l), $\widehat{\Lambda_0}$ is the estimated baseline cumulative hazard function and δ_l is a censoring indicator.

Hence, as standardized score converges to a mean-zero Gaussian process, the resulting p-values are lawful asymptotically irrespective of the factor structure. Then,

given the above empirical process, the PH assumption for the p factor can be evaluated by plotting a dozen of simulated $\widetilde{U}_p^*(t)$ on the same graph as the observed $\widetilde{U}_p^*(t)$, thus proving whether the observed scores fit the null distribution samples.

Furthermore, given equation 3.22, this graphical methods can be improved by Kolmogorov-type supremum test and the test statistic is given by

$$Sup||\widetilde{U}_m^*(\widehat{\beta}, t)|| \tag{3.23}$$

denoting the standardized empirical score process for the p^{th} component of Z. Liu (2012) stated that such test score is consistent against the alternative of non-proportionality hazards, in which the impact of at least one factor are not time independent.

Given the value of the significance level, the decision can be made on whether or not the PH assumption in Cox model is valid. That is, if the observed p-value of the Kolmogorov-type supremum test for PH assumption is less than the significance level on the given factor, that is, a non significance p-value for the supremum test suggest proportionality (Wong et al., 2011). It can be concluded that the proportionality assumption for that factor is invalid. Hence, factors with p-values greater than the significance level shows that the proportionality assumption is valid. Gharibvand et al. (2008) further indicated that such non-significant p-values shows that the is no relationship between time and residuals.

3.4.9 Parametric Survival Models

Naturally, any distribution of non-negative random variables could be used to describe durations. The distributions to be discussed here are all continuous. Throughout the literature on survival analysis, certain parametric distributions have been used repeatedly such as exponential and Weibull. These distributions have closed form expressions for survival and hazard functions. Such models are considered

as a means of increasing the precision in the estimation of small tail probability (Pollock et al., 1989).

3.4.10 The Accelerated Failure Time (AFT) Model

Although parametric PH model are very applicable to analyse survival data, there are probability distribution for survival time that can be used with these models. Therefore, the AFT model, where-in a particular form of the survival distribution is assumed, is an alternative to Cox PH model for the analysis of survival time data. AFT model means that the survival function of a patient with factor X at time t is the same as survival function of a patient with a baseline survival function at time $e^{\theta|X}t$, where $\theta|$ is the vector of the regression coefficients. That is, if X_i for $i = 1, 2, ..., p$ denotes p^{th} independent variables for each individual in the study, then, the AFT model assumes that

$$S(t|x) = S_0[\frac{t}{\eta(x)}] \tag{3.24}$$

where S_0 is a baseline survival function and $\eta(x) = e^{\theta|X} = e^{\theta_1 x_1 + \theta_2 x_2 + ... + \theta_p x_p}$ is called the accelerator factor (Constant factor). Under AFT model, the covariate effects are assumed to be constant and multiplicative on time scale, showing that the factor impacts on survival by accelerator factor (Jiezhi, 2009). From equation 3.24, it is seen that factors act multiplicative on time so that their effect is to accelerate or decelerate the time to failure relative to S_0. Furthermore, it also shows how a change in a factor alters the time scale from the baseline time scale and also implies that the median time to event with factor X is the baseline median time to event divided by its accelerator factor. The hazard function for subjects or individuals with factor $X_1, X_2, ..., X_p$ under this model is given by

$$h(t|x) = [\frac{1}{\eta(x)}]h_0[\frac{t}{\eta(x)}] \tag{3.25}$$

Furthermore, an equivalent formulation of AFT model is the following linear regression for the log transformed event time denoted by

$$ln(T_i) = \theta_0 + \theta_1 X_{1i} + \theta_2 X_{2i} + ... + \theta_p X_{pi} + \sigma\varepsilon_i \tag{3.26}$$

where θ_0 is the intercept, θ_i for $i = 1, 2, ..., p$ are regression coefficients, ε_i is a random variable assumed to have a particular distribution and σ is a scale parameter. Therefore, for each distribution of ε_i, there exist a corresponding distribution of T. The survival distribution of T_i is identified by

$$\begin{aligned} S_i(t) = P(T_i > t) &= P(ln(T_i) > ln(t)) \\ &= P(\theta_0 + \theta_1 X_{1i} + \theta_2 X_{2i} + ... + \theta_p X_{pi} + \sigma\varepsilon_i > ln(t)) \\ &= P(\varepsilon_i > \frac{ln(t) - \theta_0 - \theta x}{\sigma}) \\ &= S_{\varepsilon_i}(\frac{ln(t) - \theta_0 - \theta x}{\sigma}) \end{aligned} \tag{3.27}$$

Moreover, AFT model is used to measure the direct effect of covariates on survival time instead of the hazard, as we do in Cox PH model. It also describes the relationship between survival probabilities and the set of covariates. The AFT models includes Weibull AFT model, Exponential AFT model, Log-logistic AFT model, Gamma AFT model and Log-normal AFT model. The AFT models are named for the distribution of T, rather than the distribution of ε_i or $ln(T_i)$.

As previously noted by George et al. (2014), the effect of individual factors in the AFT models is interpreted using time ratios (TR), where the ratio denotes the acceleration factor. Contrary to HR, the TR less than one means that an event is more likely to happen and TR greater than one implies that an event is less likely to occur as it means that an investigator must wait longer for an event to occur.

Exponential AFT model

The easiest AFT model is exponential where T at $x = 0$, has exponential distribution with constant hazard $e^{-\theta_0}$. This is the same as assuming that $\sigma = 1$ and ε follows a standard extreme value distribution with a density function given by $f(\varepsilon) = e^{\varepsilon - e^{\varepsilon}}$. Therefore, e^{ε} has a standard exponential distribution with a constant hazard one. Hence, from this specification, it is simple to see that the distribution of survival time T at any factor vector x have an exponential constant hazard which is independent of t, and is therefore denoted by

$$h(t|x) = e^{\theta_0 - \theta_1 x_1 - ... - \theta_p x_p} \tag{3.28}$$

So, we automatically get the PH models. Then, suppose that for a given set of factors $(x_1, ..., x_p)$, the corresponding $S(t)$ of this model is denoted by

$$S(t|x) = e^{-h(t|x)t} \tag{3.29}$$

Now, if we let $\beta_j = -\theta_j$, then, equivalently we get

$$h(t|x) = e^{\beta_0 + \beta_1 x_1 + ... + \beta_p x_p} \tag{3.30}$$

Hence, given $k = 1, ..., p$, if the value of factor x_k is increased by one unit from x_k to x_{k+1}, while holding other factor values constant, the ratio of the corresponding hazard is given by

$$\frac{h(t|x_{k+1})}{h(t|x_k)} = e^{\beta_k} \tag{3.31}$$

and e^{β_k} can be explained as the HR corresponding to a one unit increase in the factor x_k while holding other factor values constant.

Weibull AFT model

Suppose that survival time T follows a Weibull distribution with scale and shape parameters λ and γ, respectively. That is, $W(\lambda, \gamma)$. Then, from equation 3.25, the $h(t)$ for the i^{th} individual under the AFT model is given by

$$h_i(t) = [\frac{1}{\eta_i(x)}]h_0[\frac{t}{\eta_i(x)}] = [\frac{1}{\eta_i(x)}]\lambda\gamma(\frac{t}{\eta_i(x)})^{\gamma-1} = \frac{1}{[\eta_i(x)]^{\gamma}\lambda\gamma(t)^{\gamma-1}} \tag{3.32}$$

Hence, demonstrating that a Weibull distribution have the AFT property, since the i^{th} observation's survival time is

$$W(\frac{1}{[\eta_i(x)]^{\gamma}\lambda}, \gamma) \tag{3.33}$$

Then, ε_i has a Gumbel distribution, if T_i follows a Weibull distribution. Hence, the $S(t)$ of a Gumbel distribution is denoted by $S_{\varepsilon_i} = e^{-e^{\varepsilon}}$. Also, based on equation 3.27 defined as

$$S_{\varepsilon_i}(\frac{ln(t) - \theta_0 - \theta^{\prime}x}{\sigma}) \tag{3.34}$$

where $\theta^{\prime}$ is a vector of the regression coefficient and x denotes explanatory variable. Then, the Weibull AFT model representation of the $S(t)$ for the i^{th} individual is given by

$$\begin{aligned} S_i(t) &= e^{-e^{(\frac{ln(t)-\theta_0-\theta_1X_{1i}-\theta_2X_{2i}-\ldots-\theta_pX_{pi}}{\sigma})}} \\ &= e^{-e^{(\frac{-\theta_0-\theta_1X_{1i}-\theta_2X_{2i}-\ldots-\theta_pX_{pi}}{\sigma})t^{\frac{1}{\sigma}}}} \end{aligned} \tag{3.35}$$

and its respective PH representation is given by

$$S_i(t) = e^{-e^{(\beta_1X_{1i}+\beta_2X_{2i}+\ldots+\beta_pX_{pi})\lambda t^{\gamma}}} \tag{3.36}$$

Therefore, by comparing formulas 3.35 and 3.36, we observe that the parameters γ, β_j and λ in the PH model can easily be expressed by the parameters θ_0, α_j and σ in the AFT model as follows: $\lambda = e^{-\frac{\theta_0}{\sigma}}$, $\gamma = \frac{1}{\sigma}$ and $\beta_j = \frac{-\alpha_j}{\sigma}$. Furthermore, using

the following equation

$$h(t) = \frac{f(t)}{S(t)} = \frac{-dlnS(t)}{dt} \tag{3.37}$$

The AFT representation of $h(t)$ for a Weibull model is given by

$$h_i(t) = \frac{1}{\sigma} t^{\frac{1}{\sigma}-1} e^{(\frac{-\theta_0-\theta_1 X_{1i}-\theta_2 X_{2i}-\ldots-\theta_p X_{pi}}{\sigma})} \tag{3.38}$$

Moreover, the median survival time of a Weibull AFT model for the i^{th} observation is given by

$$t_i(50) = e^{[\sigma ln(ln2)+\theta_0+\theta' x_i]} \tag{3.39}$$

Log-logistic AFT model

If the survival times have a log-logistic distribution with two parameters α and v. Then, from equation 3.25, the $h(t)$ for the i^{th} observation under the AFT model is defined as

$$h_i(t) = [\frac{1}{\eta_i(x)}]h_0[\frac{t}{\eta_i(x)}] = \frac{e^{\alpha} v[\frac{t}{\eta_i(x)}]^{v-1}}{\eta_i(x)\{1+e^{\alpha}[\frac{t}{\eta_i(x)}]^{v}\}} = \frac{e^{\alpha - v ln \eta_i(x)} v t^{v-1}}{1+e^{\alpha - v ln \eta_i(x)} t^{v}} \tag{3.40}$$

Hence, demonstrating that the log-logistic distribution have the AFT property since the i^{th} observation's survival time has a log-logistic distribution with the parameters $\alpha - v\ ln\ [\eta_i(x)]$ and v. Suppose, if the baseline $S(t)$ with unknown parameters α and v is defined by $S_0(t) = (1+e^{\alpha} t^{v})^{-1}$. Then, the baseline odds of surviving pass time t are given as

$$\frac{S_0(t)}{1-S_0(t)} = e^{-\alpha} t^{-v} \tag{3.41}$$

Also, the i^{th} observation's survival time has a log-logistic distribution which is defined as

$$S_i(t) = \frac{1}{1+e^{\alpha - v ln[\eta_i(x)]} t^{v}} \tag{3.42}$$

Therefore, the log-logistic distribution is a proportional odds model, since it have

the odds property. That is, the odds of surviving pass time t and is defined by

$$\frac{S_i(t)}{1-S_i(t)} = e^{ln[\eta_i(x)]-\alpha}t^{-v} \tag{3.43}$$

In other words, the odds of an observation surviving pass t are indicated as

$$\frac{S_i(t)}{1-S_i(t)} = e^{(\beta_1 X_{1i}+\beta_2 X_{2i}+\ldots+\beta_p X_{pi})}\frac{S_0(t)}{1-S_0(t)} \tag{3.44}$$

Then, ε_i has a logistic distribution if T_i follows a log-logistic distribution and the $S(t)$ of a logistic distribution is denoted by $S_{\varepsilon_i} = \frac{1}{1+e^{\varepsilon}}$. Also, based on equation 3.27 defined as $S_{\varepsilon_i}(\frac{ln(t)-\theta_0-\theta^{|}x}{\sigma})$. The log-logistic AFT model representation of the $S(t)$ for the i^{th} observation is given by

$$S_i(t) = \frac{1}{t^{\frac{1}{\sigma}}e^{(\frac{-\theta_0-\theta_1 X_{1i}-\theta_2 X_{2i}-\ldots-\theta_p X_{pi}}{\sigma})}+1} \tag{3.45}$$

Therefore, by comparing formulas 3.42 and 3.45, we observe that the parameters $\alpha = \frac{\theta_0}{\sigma}$, and $v = \frac{1}{\sigma}$. Then, the AFT representation of $h(t)$ for the a log-logistic model, according to the relationship of $S(t)$ and $h(t)$ for the i^{th} observation is given by

$$h_i(t) = \frac{1}{\sigma t[1+t^{\frac{-1}{\sigma}}e^{(\frac{-\theta_0-\theta_1 X_{1i}-\theta_2 X_{2i}-\ldots-\theta_p X_{pi}}{\sigma})}]} \tag{3.46}$$

Furthermore, the median survival time of a log-logistic AFT model for the i^{th} observation is given by

$$t_i(50) = e^{(\theta_0+\theta^{|}x_i)} \tag{3.47}$$

Gamma AFT model

The pdf of the generalized gamma distribution with the following parameters α, λ and the shape parameter γ is given by

$$f(t) = \frac{\alpha\lambda^{\alpha\gamma}}{r(\gamma)} t^{\alpha\gamma-1} e^{-(\lambda t)^{\alpha}} \tag{3.48}$$

where t, γ, λ and α are all positive. The $h(t)$ and $S(t)$ does not have a closed form for the generalized gamma distribution. Therefore, if $\gamma = 1$, the generalized gamma distribution becomes a Weibull distribution, and if $\gamma = \alpha = 1$, it becomes an exponential distribution. Hence, the Weibull and exponential distribution are both special cases of generalized gamma distribution.

3.4.11 Comparing parametric models

The are two ways of comparing parametric models. The Akaike information criterion (AIC), which can be used when the parametric models are not nested and the likelihood ratio (LR) test, which can be used to test for the significant deviation from more parsimonious model when the parametric models are nested. The current study apply the AIC, which is used to measures the goodness of fit of an estimated statistical method. Therefore, the AIC is defined as

$$AIC = -2l + 2(K + P) \tag{3.49}$$

where l is the log-likelihood, K is the total number of covariates in the model and P is the number of parameters with $P = 1$ indicating exponential and $P = 2$ for Weibull, log-logistic and gamma distributions. The model with smallest AIC value is considered a better fit.

3.4.12 Survival Analysis Residual

Adequacy of the fitted survival model needs to be assessed (Fitrianto and Jiin, 2013). That is, if the model is used inappropriately, it may give misleading conclusions. Hence, checking if a given model is an appropriate representation of the data is very important. The study will use residuals as tools to assess the adequacy of the model. Therefore, the following set of residuals will be used, namely Cox-Snell, Martingale and deviance residuals.

Cox-Snell Residuals

The common practice of checking the overall goodness of fit in survival models is Cox-Snell residuals (Ansin, 2015). This type of residuals was introduced by Cox and Snell in a situation more general than the usual linear hypothesis (Loynes, 1969). Cox-Snell residual is defined by

$$r_{CS_i} = e^{\widehat{\beta}^{\prime} X_i \times \widehat{H_0}(t_i)} \tag{3.50}$$

where $\widehat{H_0}(t_i)$ is an estimated cumulative hazard function (estimated integrated baseline hazard) for the fitted Cox PH model. Furthermore, Cox-Snell residual to be used for assessing the model fit of the i^{th} individual in parametric models is given by

$$r_{CS_i} = -log\widehat{S}(t_i) \tag{3.51}$$

which is the log of the estimated survival time. The Cox-Snell residuals of parametric models consider in this study are given in Table 3.2.

Table 3.2: Cox-Snell residuals of Exponential, Weibull, Log-logistic and Gamma models

Exponential	**Weibull**	**Log-logistic**	**Gamma**
$r_{CS_i} = e^{(\widehat{\beta}^{\prime} X_i)\widehat{\lambda} t}$	$r_{CS_i} = e^{(\widehat{\beta}^{\prime} X_i)\widehat{\lambda} t^{\widehat{\gamma}}}$	$r_{CS_i} = log[\frac{1}{1+e^{(\widehat{\beta}^{\prime} X_i)\widehat{\lambda} t^{\widehat{\gamma}}}}]$	$r_{CS_i} = e^{(\widehat{\beta}^{\prime} X_i)(\widehat{\lambda} t)^{\alpha}}$

Moreover, Cox and Snell (1971) demonstrated the importance of calculating a test statistics and a function of the residuals when residuals are used to assess the assumed model adequacy. Cox Snell residuals approximately follows a unit exponential distribution when the model is exactly correct. This means that r_{CS_i} are expected to have the mean of one.

Martingale and Deviance Residuals

Martingale residuals ($M_i(t)$) are slight modification of Cox Snell residuals and they are important in evaluating the functional form of a covariate. $M_i(t)$ are stochastic component and are defined by

$$M_i(t) = \delta_i(t) - e^{\widehat{\beta}' X_i \times \widehat{H_0}(t_i)} \tag{3.52}$$

where $\delta_i(t)$ denotes the number of observed events that occurs at each failure time t. However, the plots of $M_i(t)$ are difficult to interpret since they are non symmetrically distributed around zero (Ansin, 2015). Furthermore, $M_i(t)$ for censored observations take negative values.

For the Cox model, the deviance residuals r_{D_i} defined by

$$r_{D_i} = sgn(M_i(t))[-2(M_i(t) + \delta_i log(\delta_i - M_i(t)))]^{\frac{1}{2}} \tag{3.53}$$

are more symmetrically distributed around zero and therefore its plots against events time are easy to interpret, since the good model fit is demonstrated by a random scatter around zero and could be used to check outliers. The sgn ensure that the r_{D_i} has the same sign as the $M_i(t)$. The log function inflates $M_i(t)$ close to one and the square root contracts the large negative values. The importance of r_{D_i} is to assess the model prediction for individuals subjects.

3.5 Test of Significance

The threshold for the significance of a hypothesis test is the p-value at 0.05.

Chapter 4

Results

4.1 Introduction

In this chapter, the analysis of time-to-first peritonitis data of 159 kidney patients who were on PD obtained from Pietersburg Provincial Hospital and the findings of the study are presented. This Chapter starts with some survival data analysis tables and conclude with more detailed analysis of factors that contribute to the development of first peritonitis episode on kidney patients who are undergoing PD.

Before statistical survival data analysis, some important definitions need to be clarified. Firstly, the event recorded in the current study data set is time-to-first peritonitis. Peritonitis is an infection related to PD failure in patients who are on PD. Secondly, time-to-first peritonitis is recorded as the period starting from the baseline to the time that an investigation ends, death or time-to-drop out. The PD failure unrelated to time-to-first peritonitis is treated as censoring.

Censoring in the current study is defined as cases where at the end of the study

investigation or the follow-up period, the peritonitis will probably not have occurred for all patients. We do not know when or whether such patients will experience peritonitis, but we know that they were peritonitis free by the end of the observation period. Again, the study will censor patients who are lost to follow-up during the study or may have experienced a competing event such as death as a result of which further follow-up was impossible. Moreover, the transfer of patients to HD for reasons other than peritonitis and finally the transfer of patient from PKDC to other hospitals. Therefore, if one of the above mentioned censoring definitions happens to an observation, the patients will be treated as censored. Overall, there are 96 observations who developed time-to-first peritonitis and 63 who were censored observations. Observations with time-to-first peritonitis are indicated by value 1 and censored observations are indicated by value 0 as shown in Table 4.1.

Table 4.1: Total number of patients with time-to-first peritonitis (event) and without time-to-first peritonitis (censored)

Censoring indicator	**Frequency(f_i)**	**Percentage**	**Cumulative f_i**
1	96	60.3774%	96
0	63	39.6226%	159

The goal of the current study as outlined in Chapter 1, is to analyse time-to-first peritonitis data using survival analysis techniques and associated risk factors of first peritonitis episode. In other words, this chapter attempts to show how the survival function is influenced by factors such as social, biological measurements, socio-economic and demographic. The chapter will also compare the results obtained using multivariate Cox PH model and AFT model techniques. In this chapter, KM method is used to calculate the non-parametric estimates of survival distribution function among variables. This chapter will further evaluate the significance difference among variables using log-rank test.

Cox PH model is also applied to test the time dependence of variables and generate survival function plots for comparison between groups or assessing PH assumption. Moreover, generating residual plots for model diagnostic to validate if

the PH assumption is not violated using deviance plots. At last, AFT parametric modeling approach with exponential, Weibull, log-logistic and gamma distributions, will be applied to measure the direct effect of covariates on time-to-first peritonitis. To evaluate the performance of the AFT model fit in the analysis, AIC and Cox-Snell residuals are used. Furthermore, the comparison of the best fitting model between Cox PH and AFT models is done using AIC.

4.2 Descriptive statistics and KM method

The descriptive statistics are used to provide detailed information about the distribution of the factors. We obtained the categorical and baseline characteristics of 159 patients using descriptive statistics as shown in Table 4.2 and Table 4.3, respectively. Table 4.2 provides time-to-first peritonitis frequency (%), the median time-to-first peritonitis (months) and the p-values obtained by log-rank test. Table 4.3 provides the mean, standard deviations and the median for peritonitis-free group and first peritonitis group as well as the p-values for each continuous factor.

Categorical variables evaluated are gram, water, electricity, employment, dwelling, seasons, sex, race, education, house type, anuric at baseline, dialysate, diastolic-dysfn, LVH, PHT and cause of ESRD. The continuous variables are age at baseline, alk-phosp, BMI, cholesterol, CaCorr, DBP, distance to dialysis center, ferritin, Hb, number of rooms, number of people, PTH, Pi, Product, SBP, Trans Sat, eGFR-MDRD, albumin, duration on PD and number of anti-HTN drugs. For categorical factors, the KM curves are generated to assess the survival time for the incidence of time-to-first peritonitis for each subcategory. Furthermore, the log-rank test is used to assess the significance of the effect of factors by comparing survival time among different groups.

4.2.1 Categorical biological and social factors

Table 4.2: Descriptive statistics for all categorical biological and social factors

Covariate	Category	N	First PX Episode in frequency(%)	Median time to PX in Months		KM p-value
Gram	Gram-positive	15	15 (100%)	08.55	[3.72,16.14]	0.0144
	Gram-negative	22	22 (100%)	04.49	[1.71,6.77]	
Anuric at baseline	Anuric	24	10 (41.7%)	23.97	[6.54,.]	0.3292
	Non-Anuric	133	84 (63.2%)	14.07	[8.78,18.74]	
Dialysate	Adcock-Icod	07	05 (71.4%)	14.14	[3.88,.]	
	Bicavera	36	24 (66.7%)	12.13	[6.77,24.92]	0.8862
	Dianeal-PD-Sol	90	48 (53.3%)	16.83	[9.34,25.05]	
	Stay-safe	26	19 (73.1%)	11.05	[5.33,16.14]	
Dystolic-dysfn	No	78	51 (65.4%)	12.03	[8.78,22.16]	0.3850
	Yes	40	22 (55.0%)	20.42	[12.13,25.05]	
LVH	No	56	37 (66.1%)	12.03	[8.02,23.05]	0.3399
	Yes	62	36 (58.1%)	17.92	[8.94,26.24]	
PHT	No	107	67 (62.6%)	14.14	[9.9,22.16]	0.4599
	Yes	11	06 (54.6%)	42.38	[1.64,.]	
Season	Autumn	43	25 (58.1%)	16.14	[8.02,24.92]	
	Spring	42	22 (52.4%)	20.42	[9.9,48.69]	0.2139
	Summer	35	24 (68.6%)	13.81	[6.28,22.16]	
	Winter	39	25 (64.1%)	06.18	[3.65,14.14]	
Sex	Female	71	51 (65.4%)	341	[199,512]	0.2584
	Male	81	45 (55.6%)	483	[339,762]	
Race	Black	147	92 (62.6%)	13.81	[9.34,17.92]	
	Indian	2	01 (50.0%)	05.85	[.,.]	0.2832
	White	10	03 (30.0%)	00.00	[0.33,.]	
Water	No	55	45 (81.8%)	09.47	[5.49,15.95]	0.0047
	Yes	101	49 (48.5%)	22.16	[11.21,27.81]	
Electricity	No	6	05 (83.3%)	03.25	[2.56,15.88]	0.0092
	Yes	150	89 (59.3%)	14.76	[11.15,22.16]	
Employment	Employed	30	11 (36.7%)	00.00	[8.78,.]	
	Scholar	16	07 (43.8%)	09.89	[5.85,.]	0.0341
	Unemployed	108	77 (71.3%)	11.24	[6.58,16.14]	
Dwelling	Rural	140	92 (65.7%)	12.03	[8.78,16.31]	0.0068
	Urban	18	04 (22.2%)	00.00	[11.31,.]	
House type	Formal	150	90 (60.0%)	14.14	[9.9,20.42]	0.1907
	Informal	6	04 (66.7%)	04.14	[0.03,.]	
Education	No formal education	6	04 (66.7%)	09.91	[1.58,15.88]	
	Primary	56	29 (51.8%)	11.31	[6.71,.]	0.5231
	Secondary	86	58 (67.4%)	14.76	[9.9,20.42]	
	Tertiary	11	05 (45.5%)	13.81	[1.12,.]	

Abbreviations: Dysfn: Dysfunction, LVH: Left ventricular hypertrophy, PHT: Pulmonary Hypertension, Season: Season of the catheter insertion, PD-Sol: Peritoneal dialysis solution

The frequency, the median and the p-values (log rank test) of time-to-first peritonitis episode for categorical social, demographic, and biological factors are presented in Table 4.2. Table 4.2 helps to check categorical factors which might be considered significant prognostic importance. Therefore, the factor gram (p=0.0144),

water (p=0.0047), electricity (p=0.0092), employment (p=0.0341), and dwelling (p=0.0068) were observed to be of significant prognostic importance towards the development of time-to-first peritonitis, since their p-values when log-rank test is applied are below the threshold of 0.05. However, anuric at baseline, dialysate, diastolic dysfunction, LVH, PHT, season, sex, race, house type, education were found not to be significant, with the p-values of greater than 0.05, indicating that the survival of PD patients could be independent of these factors.

However, out of 159 kidney patients who are on PD, 96 of them had peritonitis. There were slightly more males (81) than females (78). But, the study found that more females 51 (65.4%) are more likely to develop time-to-first peritonitis episode than males 45 (55.6%). A majority of 140 patients are from rural areas with 92 (65.7%) of them experiencing first episode of peritonitis compared to 18 patients coming from urban areas with only 4 (22.2%) of them experiencing first episode of peritonitis.

Furthermore, most patients had water (101) and few did not have water (55). However, out of those without water, 45 (81.8%) had experienced first episode of peritonitis, with only 49 (48.5%) of those with water experiencing first episode of peritonitis. In terms of electricity, majority of patients had electricity (150) and few (6) of them did not have. Therefore, the study found out that 5 (83.3%) of the patients with no electricity had experienced first episode of peritonitis, with 89 (59.3%) of those with electricity experiencing first episode of peritonitis. It is also observed that in unemployed patients (108), 77 (71.3%) of them had developed first episode of peritonitis, as compared to scholar patients (16), with 7(43.8%) of them experiencing first episode of peritonitis. As for employed patients (30), only 11(36.7%) patients had peritonitis. Causative organisms of first peritonitis episode were due to gram positive organisms (15), with all of them experiencing peritonitis and gram negative organisms (22), with all of them experiencing peritonitis. We also observed that majority of patients without dystolic dysfunction , LVH, PHT, and having informal housing as well as being non-anuric had experienced higher first episode of

peritonitis.

To validate if indeed these factors should or should not be considered as potential risk factors for first peritonitis episode in our study population. Univariate Cox PH model will be applied to further assess and validate the significant impact of all these factors and thereby including those which were observed to be below the threshold of 0.05 in the final Cox PH model.

4.2.2 Continuous biological and social factors measured at baseline

Table 4.3: Descriptive statistics for continuous biological and social factors measured at baseline

Covariate	**Peritonitis free group**		**First peritonitis group**		**P-value**
	N	[mean±std dev, median]	N	[mean±std dev, median]	
Age at baseline	63	[33.68±11.65, 34.00]	96	[34.42±11.86, 36.50]	0.7012
Weight	56	[65.24±17.12, 65.20]	86	[64.89±17.30, 62.00]	0.9059
Alk-Phosp (mmol/L)	40	[129.05±125.36, 87.00]	63	[105.29±83.84, 87.00]	0.2942
BMI (kg/m^2)	57	[24.45±4.83, 23.90]	92	[24.15±4.99, 24.00]	0.7243
Cholestorol ($mmol/L$)	36	[4.53±1.35, 4.30]	56	[4.49±1.37, 4.39]	0.8856
CaCorr ($mmol/L$)	58	[2.28±0.27, 2.34]	92	[2.31±0.32, 2.35]	0.5382
DBP ($mmhg$)	58	[88.57±20.19, 86.50]	83	[88.53±19.14, 85.00]	0.9908
Ferritin (ug/l)	57	[428.75±351.74, 300.00]	86	[467.35±432.10, 358.50]	0.5750
Hb (g/dL)	63	[9.05±2.19, 9.10]	96	[8.52±1.98, 8.40]	0.1138
number of anti-HTN drugs	48	[2.29±1.20, 2.00]	85	[2.44±1.18, 2.00]	0.5042
PTH (ng/L)	44	[461.59±573.23, 251.5]	65	[419.54±353.28, 329.00]	0.6657
Pi ($mmol/L$)	55	[1.82±0.74, 1.68]	91	[1.72±0.98, 1.48]	0.4767
SBP ($mmHg$)	58	[142.64±25.49, 139.5]	83	[145.36±30.29, 142.00]	0.5764
Trans-Sat (%)	45	[27.13±14.77, 23.00]	68	[34.26±22.37, 29.00]	0.0436
eGFR MDRD ($mL/min/1.73m^2$)	63	[6.42±3.65, 5.00]	96	[6.76±3.28, 6.00]	0.5394
Duration on PD	63	[18.87±14.04, 17.00]	96	[26.67±20.81, 21.00]	0.0054
Albumin	60	[30.28±6.66, 30.50]	92	[30.03±5.99, 30.00]	0.8097
Distance to dialysis center (km)	63	[111.81±76.69, 112.00]	96	[129.27±71.14, 126.50]	0.1442
Number of rooms	61	[4.90±1.72, 5.00]	91	[5.19±2.41, 5.00]	0.3964
Number of people	58	[4.52±2.10, 4.00]	91	[5.05±2.41, 5.00]	0.1655

Abbreviations: Alk Phos: Alkaline phosphatase, BMI: Body mass index, CaCorr: calcium corrected, DBP: Diastolic blood pressure, Hb: hemoglobin, Anti-HTN: anti hypertensive, PTH: Parathyroid hormone, Pi: phosphorus, SBP: systolic blood pressure, Trans-Sat: saturation of transferrin, eGFR MDRD: glomerular filtration rate according to MDRD formula, PD: Per-itoneal dialysis

Table 4.3 provides continuous baseline characteristics for patients who had first peritonitis episode and those who did not during the follow-up period. The purpose of Table 4.3 is to compare the group of patients who had first peritonitis episode and peritonitis-free patients group and identify continuous prognostic factors that

may contribute to the development of first peritonitis episode. Therefore, factors tran-sat (p=0.0436) and duration on PD (p=0.0054) based on t-test were observed to be statistically significant. However, age baseline, alk phosp, BMI, cholesterol, CaCorr, DBP, distance KM, ferritin, Hb, number of rooms, number of people, PTH, Pi, Product, SBP, eGFR MDRD, and albumin were observed not to be significant, indicating that survival of patients could be independent of these factors. Therefore, to validate these observations, univariate Cox PH model is performed.

4.2.3 KM survival curves for the categorical factors

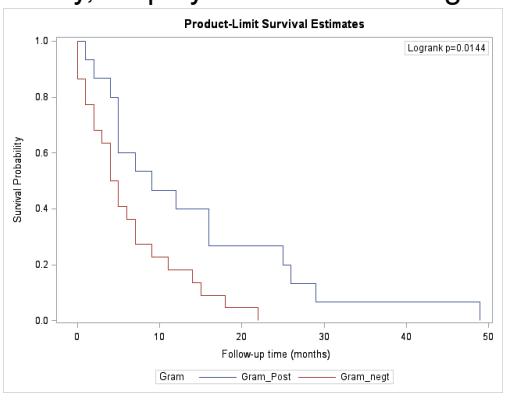
The following KM curves are for factors which were found to be significant in Table 4.2. These curves provide the first insight of the shape of the survival function for the incidence of time-to-first peritonitis among each treatment category.

The KM survival curves in Figure 4.1 (a) clearly shows that there is a significant difference between patients with gram negative organisms and those with gram positive organisms. Those with gram negative experience first peritonitis more than those with gram positive organisms. The KM survival curves also confirmed the results showed by the log-rank test (p=0.0144) in Table 4.2. Referring to KM curve in Figure 4.1 (b), survival is dependent on water, with patients who have no water showing low survival than those with water. This was also confirmed by the log-rank test (p=0.0047).

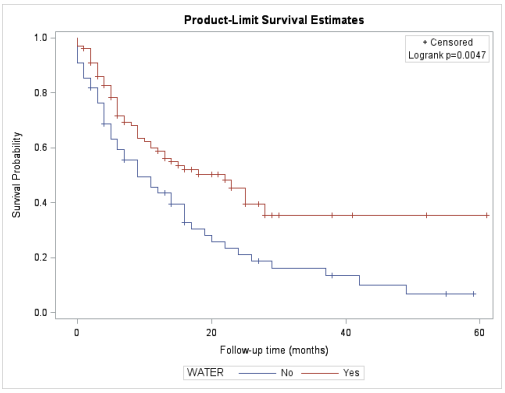
Figure 4.1 (c) indicates that at the initial stage of the treatment, no significant difference is observed between patients with electricity and those who do not. However, at the later stage, the curves show high survival of patients who have electricity. Although clear significance of electricity in predicting survival was not well observed, since the two curves do cross at the early stage of the treatment, we can conclude that electricity show a significant difference in predicting the survival among these two groups of patients, since a high significance when using log-rank test (p=0.0092) was observed.

Using log-rank test, we found that the three groups of employment are significant with p-value of 0.0341. But the log-rank test was unable to tell which of the two groups contribute to this difference. However, Figure 4.1 (d), shows that there is a clear significant difference between patients who are employed and those who are unemployed. Patients who are employed have higher survival as compared to those who are unemployed. No significant difference was observed between employment and scholar as well as unemployment and scholar. From Figure 4.1 (e), the KM survival curves show that survival is dependent on dwelling, with patients staying in rural areas having lower survival than those staying in urban areas. Similar results were found using log-rank test (p=0.0068)

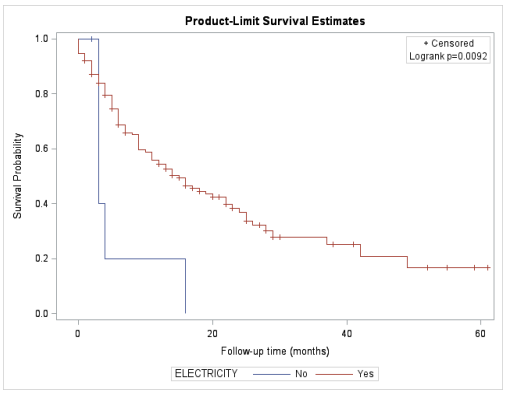
Figure 4.1: The KM survival distribution of treatment groups for factors gram, water, electricity, employment and dwelling.



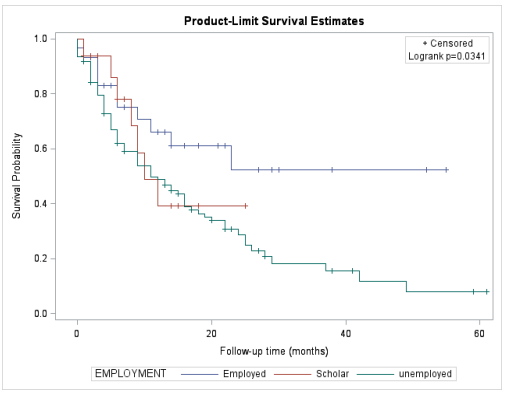
(a)



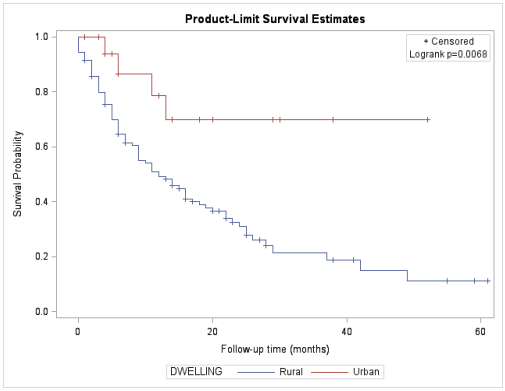
(b)



(c)



(d)



(e)

4.3 Cox PH model selection

To determine factors which will be included in the final multivariate Cox PH model, the univariate Cox PH model analysis is applied first to identify all risk factors of time-to-first peritonitis before proceeding to multivariate Cox PH model. The threshold for model selection to maintain both continuous and categorical risk factors from univariate PH model and include them in the final Cox PH model is p-value of atmost 0.05. However, factors with high levels of multicollinearity will only be assessed in the univariate models. Although, they may be found to be significantly associated with time-to-first peritonitis, one of them will be excluded in the multivariate models.

4.3.1 Univariate Cox PH model analysis for categorical factors

The univariate Cox PH model analysis of each factor using Wald test gives a preliminary idea of which factors could have an impact towards the development of time-to-first peritonitis episode. We apply this model to identify the impact of each risk factors before proceeding to multivariate Cox PH model. The univariate Cox PH model for all categorical factors fitted to this time-to-first peritonitis data set is obtained as indicated in Table 4.4. The table shows the estimates, HR, 95% HR confidence limits and the p-values for each factor.

Table 4.4, reveals that factor gram, season, water, electricity, employment and dwelling were significantly associated with time-to-first peritonitis episode, since gram-negative (HR=2.366, 95% CI=1.129 to 4.962, p=0.0226), winter season (HR= 1.814, 95% CI=1.019 to 3.228, p=0.0428), unavailability of water (HR=1.765, 95% CI=1.174 to 2.652, p=0.0063), unavailability of electricity (HR=3.094, 95% CI=1.238 to 7.734, p=0.0157), unemployment (HR=2.211, 95% CI=1.174 to 4.165, p=0.0148) and rural dwellers (HR=3.564, 95% CI=1.308 to 9.710, p=0.0130) have p-values below the threshold of 0.05, respectively. But sex, race, dialysate, diastolic dys-

function, LVH,PHT, causes of ESRD, house type and education were not statistically significant, which suggest that this factors are not associated with time-to-first peritonitis in this study population.

With regards to HR of statistically significant factors, univariate analysis shows that the rate of experiencing time-to-first peritonitis for gram-negative patients is 2.366 times higher than gram-positive patients, while holding other factors constant. For the season of catheter insertion, the rate of experiencing time-to-first peritonitis for patients who inserted their catheter in winter is 1.814 times higher as compared to patients who inserted their catheter in spring while holding other factors constant. However, for autumn (HR=1.285, 95% CI=0.723 to 2.282, p=0.3931) and summer (HR=1.421, 95% CI=0.796 to 2.538, p=0.2345), the hazard ratio is not statistically significant. The model also revealed that the rate of time-to-first peritonitis for patients with no water is 1.765 times higher than patients who have water, while holding all other factors constant.

Moreover, the model also showed that the rate of experiencing time-to-first peritonitis for patients with no electricity is 3.094 times higher than patients with electricity, while holding all other factors constant. With respect to dwelling, the rate of experiencing time-to-first peritonitis for patients coming from rural areas is 3.564 times higher than patients coming from urban areas, while holding all other factors constant. Furthermore, the model found that the rate of experiencing time-to-first peritonitis for unemployed patients is 2.211 times higher than patients who are employed, while holding other factors constant. But, for scholar patients (HR=1.739, 95% CI=0.670 to 4.518, p=0.2258), the hazard ratio is not statistically significant.

However, for collinearity among variables, significant correlation between gram organisms and season of the catheter insertion was observed, and also correlation between employment and electricity as well as water. Therefore, gram and employment will be excluded in our multivariate analysis. Hence, only categorical factors season, water, electricity and dwelling will be entered in the multivariate Cox PH

model.

Table 4.4: Univariate Cox PH model analysis for Categorical biological and social factors

Covariate	Category	β	HR	95% LCL HR	95% UCL HR	P-value
Gram	Gram-negative	0.861	2.366	1.129	4.962	0.0226
	Gram-positive (ref)	0.000	1.000			
Anuric at baseline	Anuric	-0.319	0.727	0.377	1.404	0.3425
	Non-Anuric (ref)	0.000	1.000			
Dialysate	Adcock-Icod	-0.064	0.938	0.348	2.527	0.8991
	Bicavera	-0.158	0.854	0.467	1.560	0.6074
	Dianeal-PD-Sol	-0.209	0.812	0.473	1.392	0.4478
	Stay-safe (ref)	0.000	1.000			
Dystolic-dysfn	No	0.338	1.403	0.848	2.322	0.1880
	Yes (ref)	0.000	1.000			
LVH	No	0.220	1.246	0.786	1.977	0.3495
	Yes (ref)	0.000	1.000			
PHT	No	0.311	0.365	0.588	3.170	0.4694
	Yes (ref)	0.000	1.000			
Season	Autumn	0.250	1.285	0.723	2.282	0.3931
	Summer	0.352	1.421	0.796	2.538	0.2345
	Winter	0.596	1.814	1.019	3.228	0.0428
	Spring (ref)	0.000	1.000			
Sex	Female	0.228	1.256	0.838	1.881	0.2694
	Male (ref)	0.000	1.000			
Race	Black	0.770	2.160	0.682	6.839	0.1904
	Indian	1.522	4.581	0.471	44.533	0.1896
	White (ref)	0.000	1.000			
Water	No	0.568	1.765	1.174	2.652	0.0063
	Yes (ref)	0.000	1.000			
Electricity	No	1.129	3.094	1.238	7.734	0.0157
	Yes (ref)	0.000	1.000			
Employment	Scholar	0.553	1.739	0.670	4.518	0.2258
	Unemployed	0.794	2.211	1.174	4.165	0.0148
	Employed (ref)	0.000	1.000			
Dwelling	Rural	1.271	3.564	1.308	9.710	0.0130
	Urban (ref)	0.000	1.000			
House type	Formal	-0.646	0.524	0.191	1.435	0.2087
	Informal (ref)	0.000	1.000			
Education	No formal education	0.6696	1.953	0.522	7.307	0.3199
	Primary	0.021	1.022	0.395	2.641	0.9647
	Secondary	0.249	1.283	0.514	3.203	0.5935
	Tertiary (ref)	0.000	1.000			

Abbreviations: Dysfn: Dysfunction, LVH: Left ventricular hypertrophy, PHT: Pulmonary Hypertension, Season: Season of the catheter insertion, PD-Sol: Peritoneal dialysis solution

4.3.2 Univariate Cox PH model analysis for continuous factors

For continuous factors, univariate Cox PH model analysis is also applied. The outcomes of univariate Cox PH model for continuous factors are summarised in Table 4.5. Tran-sat (HR=1.011, 95% CI=0.1.001 to 1.022, p=0.0398) and Hb (HR=0.889, 95% CI=0.807 to 0.979, p=0.0164) were observed to be associated with higher risk of experiencing first peritonitis episode. Therefore, they will be entered in the multivariate Cox PH model, since their p-values are less than 0.05.

With regard to their HR, the HR implies that as the value of tran-sat increases by one unit, while keeping all other factors constant, the rate of experiencing time-to-first peritonitis increases by 1.1. Also, an increase in Hb concentration of a patient will decrease the the rate of experiencing time-to-first peritonitis by 0.111, while assuming that all factors are constant. However, age at baseline, weight, eGFR-MDRD, alk-phosphate, BMI, cholesterol, CaCorr, ferritin, PTH, number of anti-NTH, Pi, SBP, distance from the dialysis center, number of rooms, number of people, albumin and duration on PD were not associated with a risk of experiencing first peritonitis episode.

Table 4.5: Univariate Cox PH model analysis for Continuous biological and social factors measured at baseline

Covariate	β	HR	95% LCL HR	95% UCL HR	P-value
Age at baseline	0.0001	1.000	0.983	1.018	0.9895
Weight	-0.002	0.998	0.986	1.011	0.8142
Alk-Phosp (mmol/L)	-0.002	0.998	0.995	1.002	0.3111
BMI (kg/m^2)	-0.016	0.984	0.943	1.028	0.4744
Cholestorol ($mmol/L$)	-0.007	0.993	0.800	1.233	0.9481
CaCorr ($mmol/L$)	0.079	1.082	0.526	2.227	0.8311
DBP ($mmhg$)	0.002	1.002	0.991	1.013	0.7395
Ferritin (ug/l)	0.0004	1.000	1.000	1.001	0.1234
Hb (g/dL)	-0.118	0.889	0.807	0.979	0.0164
No anti-HTN drugs	0.037	1.038	0.861	1.251	0.6961
PTH (ng/L)	-0.0003	1.000	0.999	1.000	0.2480
Pi ($mmol/L$)	-0.053	0.948	0.736	1.222	0.6818
SBP ($mmHg$)	0.004	1.004	0.996	1.012	0.2836
Trans-Sat (%)	0.011	1.011	1.001	1.022	0.0398
eGFR MDRD ($mL/min/1.73m^2$)	-0.001	0.999	0.944	1.057	0.9737
Duration on PD	-0.011	0.989	0.976	1.002	0.1022
Albumin	-0.018	0.982	0.949	1.015	0.2819
Distance to dialysis center (km)	0.003	1.003	1.000	1.005	0.0521
Number of rooms	-0.007	0.993	0.896	1.100	0.8905
Number of people	0.018	1.018	0.935	1.108	0.6848

Abbreviations: Alk Phos: Alkaline phosphatase, BMI: Body mass index, CaCorr: calcium corrected, DBP: Diastolic blood pressure, Hb: hemoglobin, Anti-HTN:anti hypertensive, PTH: Parathyroid hormone, Pi: phosphorus, SBP: systolic blood pressure, Trans-Sat: saturation of transferrin, eGFR MDRD: glomerular filtration rate according to MDRD formula, PD: Peritoneal dialysis

4.3.3 Testing for proportional hazard assumption

The multivariate Cox PH model is constructed on the assumption that the hazards between two or more groups are proportional. Therefore, this study used the KM curves of the $(log(-log(S(t))))$ vs $log(t)$ to validate the assumption of proportionality for the selected categorical factors. This is due to the fact that KM curves are usually used for categorical factors, as well as continuous factors that may be easily grouped. The KM curves between groups should be parallel if the Cox PH assumption is met. Figure 4.2 shows curves of $(log(-log(S(t))))$ vs $log(t)$ for factor season, water, electricity and dwelling.

Figure 4.2: The KM survival distribution of treatment groups by $(log(-log(S(t))))$ vs $log(t)$ for factors season, water, electricity, and dwelling.

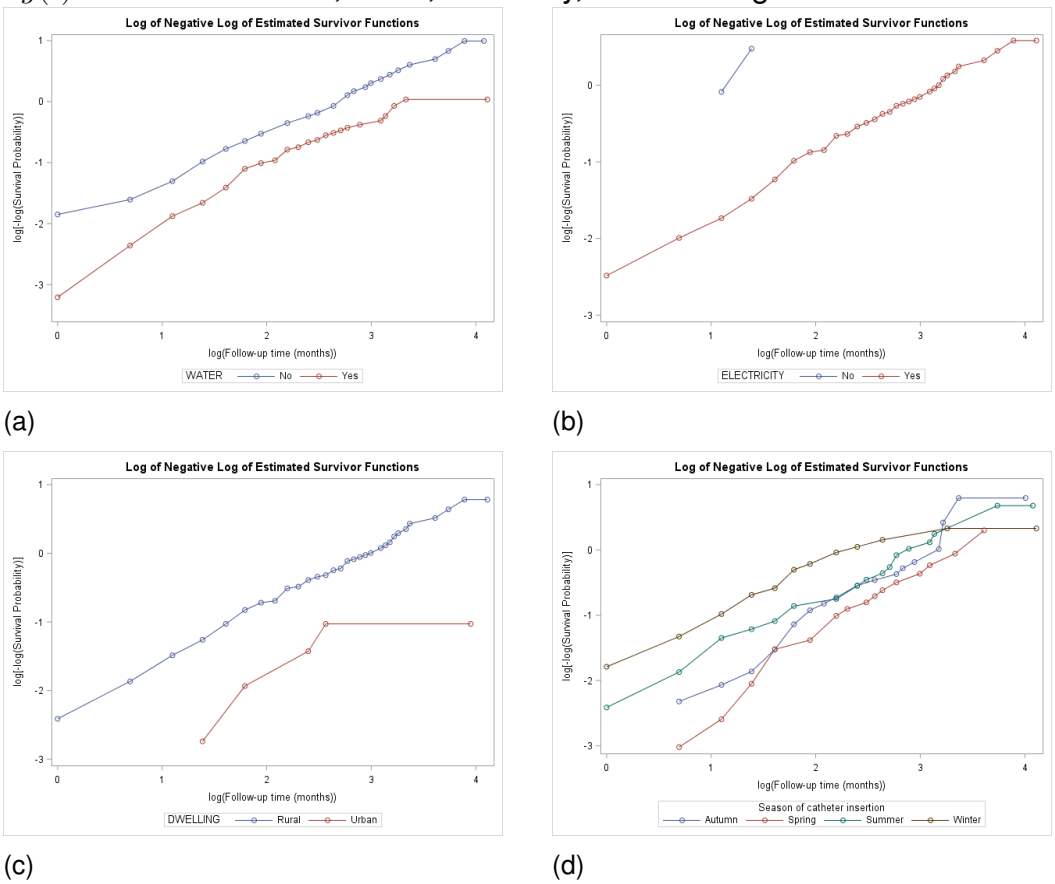


Figure 4.2 (a) shows that the KM curves of patients with water and those without water are roughly parallel, which indicates that the PH assumption is not violated. Figure 4.2 (b) also shows that the KM curves of patients with electricity and those who do not have electricity are roughly parallel, indicating no violation of Cox PH assumption. Figure 4.2 (c) similarly shows that the KM curves of patients who come from rural areas and those who come from urban areas are roughly parallel, indicating no violation of Cox PH assumption. KM curves in Figure 4.2 (d) indicate roughly parallel curves between, treatment group spring and winter, spring and autumn, spring and summer seasons, demonstrating that Cox PH assumption is not violated between patients who have catheter insertion during spring and all the seasons. However, curves for winter, autumn and summer crosses each other,

indicating a possible violation of Cox PH model.

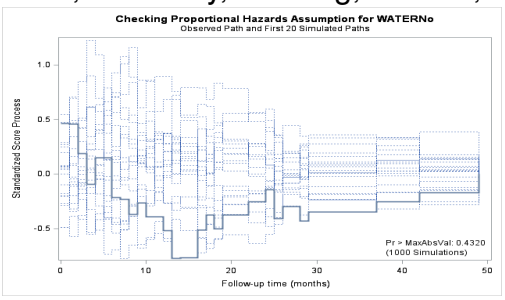
4.3.4 Validation of Cox PH model assumption

The validation of Cox PH assumption is performed using cumulative martingale residual process and supremum test. The graphical cumulative martingale residual process check is helpful in providing more confident conclusion about PH assumption and so is the supremum test.

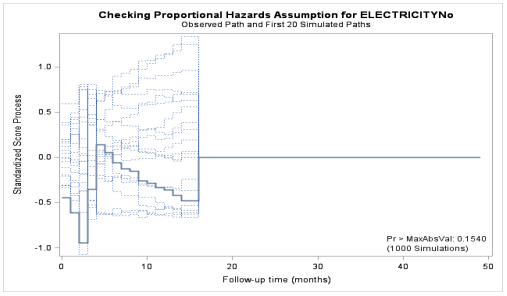
Checking proportional hazard assumption with cumulative sums of martingale-based residuals

The plots in Figures 4.3 displays graphical results of a proportional hazard assumption check for water, electricity, dwelling, season, Hb and tran-sat. Figures 4.3 (a) and (b), suggest that in an early stage of the life course, the witnessed scores are constantly below zero, thus indicating some systematic variability. However, the overall observed process tends to fluctuate randomly around zero. Therefore, we conclude that the null hypothesis of the proportionality assumption about water and electricity cannot be rejected. Moreover, Figures 4.3 (c), (g) and (h) displays plots of the observed cumulative martingale residual process for dwelling, Hb and tran-sat. According to this plots, the observed martingale residuals does not seem to fall off systematically from the null hypothesis distribution. Furthermore, the plots in Figures 4.3 (d), (e) and (f) also appear not to fall systematically from the null hypothesis, except at some survival times located in the middle. Since the level of significant is above 0.05, we can confidently say the observed martingale process for these factors does not deviate significantly from the simulated realisations. Hence, using another transform of these factors like stratifying seems unnecessary. As a result, we can have sufficient confidence to conclude that using of water, electricity, dwelling, season, Hb and tran-sat in a final Cox PH model is more appropriate given the simplicity attaching to these functional forms.

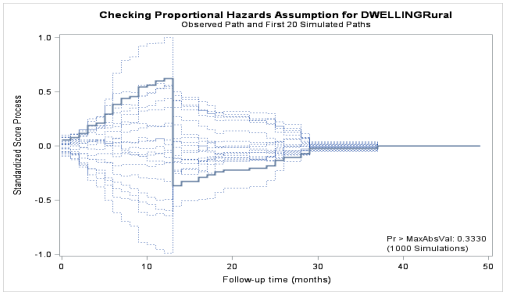
Figure 4.3: The plots displays the observed cumulative martingale residual process for water, electricity, dwelling, season, Hb and tran-sat.



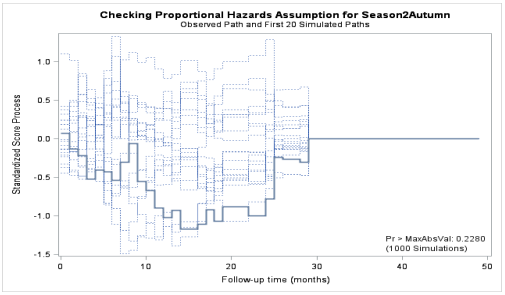
(a)



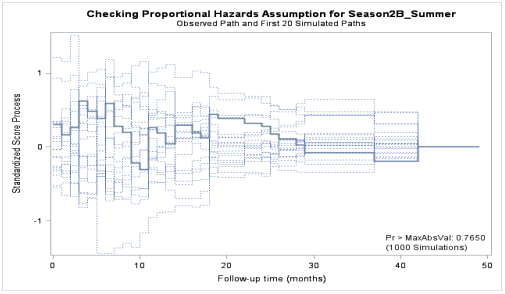
(b)



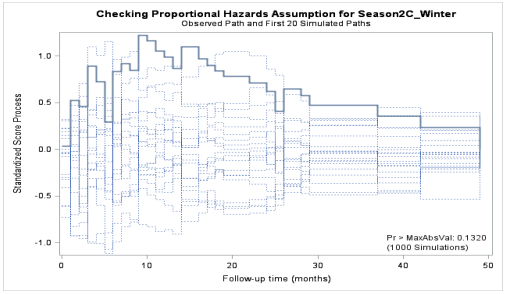
(c)



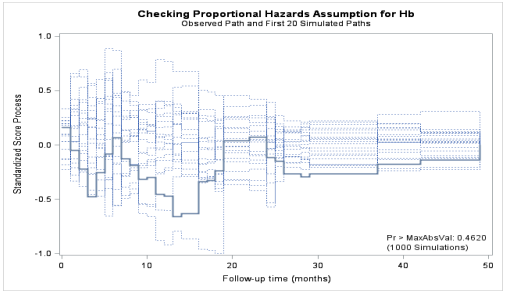
(d)



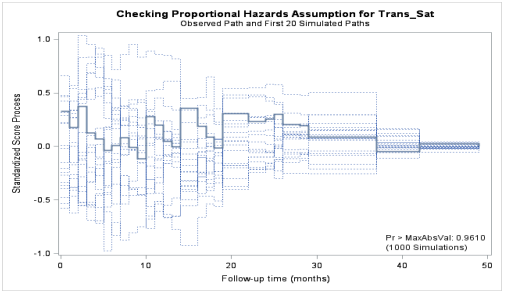
(e)



(f)



(g)



(h)

Supremum test for proportional hazard assumption

Table 4.6: Supremum test for proportional hazard assumption

Covariate	MaxAbsValue	Replications	Sead	P-value
Water (No)	0.7732	1.000	635830310	0.4320
Electricity (No)	0.9494	1.000	635830310	0.1540
Dwelling (Rural)	0.6239	1.000	635830310	0.3330
Season (Autumn)	1.1723	1.000	635830310	0.2280
Season (Summer)	0.6259	1.000	635830310	0.7650
Season (Winter)	1.2229	1.000	635830310	0.1320
Hb	0.6623	1.000	635830310	0.4620
Tran-sat	0.3727	1.000	635830310	0.9610

Abbreviations: Season: Season of the catheter insertion, Hb: haemoglobin, Trans-Sat: saturation of transferrin

The validation of Cox PH assumption can also be done by supremum test. Therefore, based on the supremum test, PH assumption is seriously violated if one of the factors have a p-value below the threshold (Gharibvand et al., 2008; Gharibvand and Fernandez, 2008). Based on Table 4.6, we can see that the p-values for all factors are more than the threshold of 0.05. Hence, we conclude that there is no serious violation of Cox PH assumption and therefore no remedial measure (stratified analysis) for correcting Cox PH assumption is needed. We are also 95% sure that there was no relationship between time and residuals.

4.3.5 Multivariate Cox PH model

The multivariate Cox PH model, shown in Table 4.7, demonstrate that categorical factors associated with an increased HR of first peritonitis episode were no water (no water versus water: HR=1.795, 95% CI=1.068 to 3.017, p=0.0272), no electricity (no electricity versus electricity: HR=3.273, 95% CI=1.107 to 9.677, p=0.0321), rural dwellers (rural versus urban: HR=8.998, 95% CI=1.202 to 67.358, p=0.0.0324) and also demonstrated that Hb (HR=0.846, 95% CI=0.753 to 0.950, p=0.0047) is the only continuous factor associated with higher risk of experiencing

first peritonitis episode.

Therefore, with regard to HR, the model revealed that, the rate of experiencing time-to-first peritonitis for patients with no water is 1.795 times higher than patients with water, while holding other factors constant. The model also shown that, the rate of experiencing time-to-first peritonitis for patients with no electricity is 3.273 times higher than patients with electricity, while holding other factors constant. Moreover, the model shown that, the rate of experiencing time-to-first peritonitis for patients coming from rural areas is 8.998 times higher than patients coming from urban areas, while holding all other factors constant.

Furthermore, it was also revealed that Hb has an estimated HR of 0.846, which implies that an increase in Hb concentration of patients will decrease the rate of experiencing time-to-first peritonitis by 0.154, while holding all the factors constant. It was also observed that, season of the inserted catheter and Tran-sat were not significant determinants of PD patients survival at 5% significant level.

Table 4.7: Results of multivariate Cox PH model

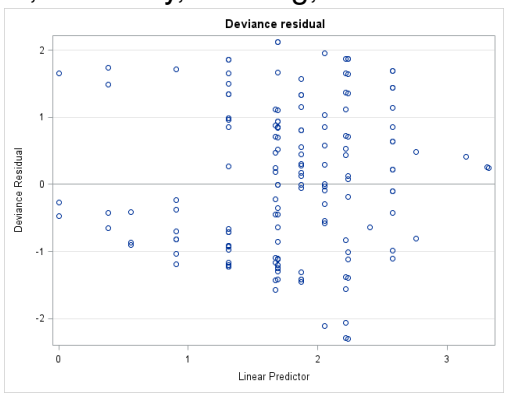
Covariate	**Category**	β	**HR**	**95% LCL HR**	**95% UCL HR**	**P-value**
Water	No	0.585	1.795	1.068	3.017	0.0272
	Yes (ref)	0.000	1.000			
Electricity	No	1.186	3.273	1.107	9.677	0.0321
	Yes (ref)	0.000	1.000			
Dwelling	Rural	2.197	8.998	1.202	67.358	0.0324
	Urban (ref)	0.000	1.000			
Season	Autumn	0.341	1.407	0.705	2.808	0.3327
	Summer	0.489	1.632	0.765	3.482	0.2055
	Winter	0.527	1.693	0.822	3.489	0.1535
	Spring (ref)	0.000	1.000			
Hb		-0.167	0.846	0.753	0.950	0.0047
Tran-sat		0.010	1.010	0.999	1.023	0.0864

Abbreviations: Ref: Reference, Season: Season of the catheter insertion, Hb: haemoglobin, Trans-Sat: saturation of transferrin

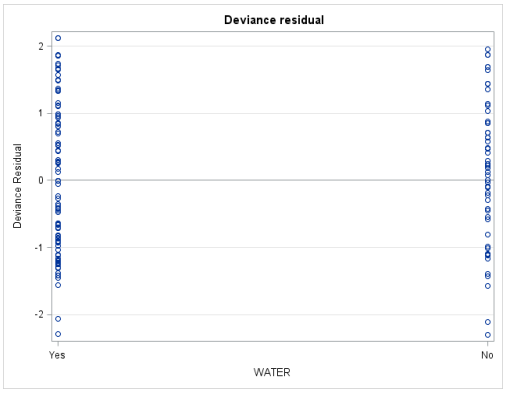
4.3.6 Model Diagnostics: Deviance residuals plots

Model diagnostic is performed by measuring deviance residual plots for all linear factors or each factor in the multivariate Cox PH model. The deviance residuals consist information about influential data and potential outliers. The patient is considered to be an outlier, if the value of the corresponding residual is not within the range of -2.5 and 2.5 and a patient with large bubble is considered as an influential observation. The deviance residuals in Figure 4.4 are positive for patients with survival times that are smaller than expected and negative for patients who had longer survival times than expected. All the deviance residual plots in Figure 4.4 shows that all the residuals converges around the baseline, and falls within the the range of -2.5 and 2.5. Therefore, there is no evidence of outliers in all the plots. The residual plots also show no evidence of influential observations. Hence, after observing the results, we can conclude that the fitting is good and there is no need to remove a particular patient from the analysis. The results also mean that there are no patients with unusual large positive deviance residual or large negative deviance residual of hospitalisation duration (PD duration) regarding their PD treatment than the expected hospitalisation months.

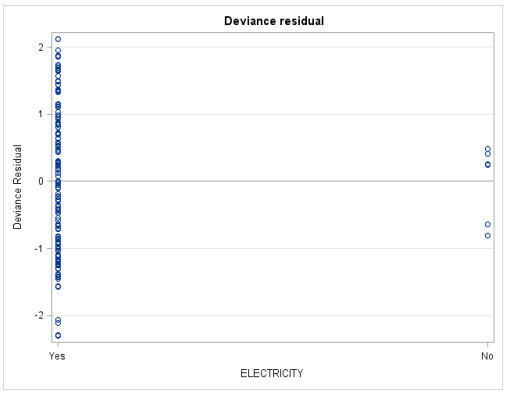
Figure 4.4: The deviance plots for the all linear factors or individual factor season, water, electricity, dwelling, Hb and TranSat.



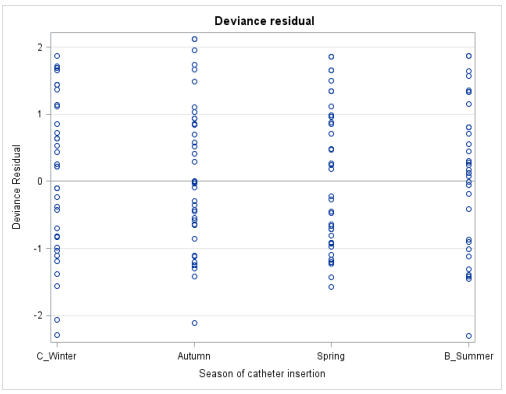
(a)



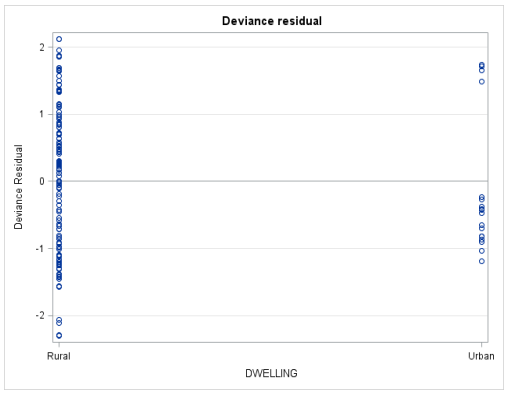
(b)



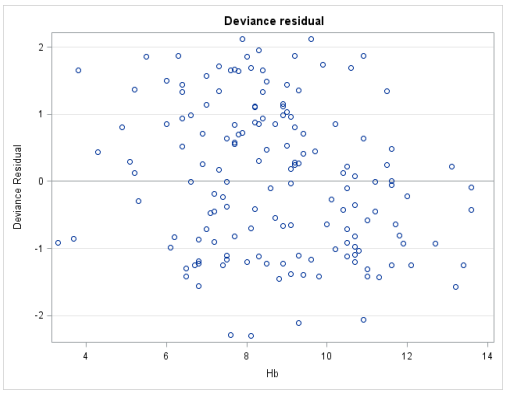
(c)



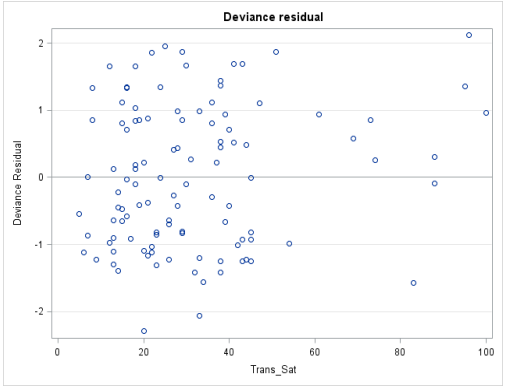
(d)



(e)



(f)



(g)

4.4 Parametric AFT models

We fit time-to-first peritonitis data set using Weibull, Exponential, Log-logistic and Gamma AFT models. The AFT models are used to assess the difference between the treatment in survival time by expressing the magnitude of impact in more accessible way. For each kind of the model, we fit univariate and multivariate AFT models for the data analysis. Univariate AFT analysis is used to identify additional significant factors to those which were selected by Cox PH model. That is, all factors which were found to be significant in any of the univariate AFT models, will collectively be combined with factors which were included in the final Cox PH model and therefore, be entered in all multivariate AFT models. Multivariate AFT models will be used to assess the effectiveness of those factors. To determine additional factors, which should be combined with factors entered in the final Cox PH model, and be used for the building of multivariate AFT models, the threshold with p-value of atmost 0.05 in the univariate AFT models will be used.

4.4.1 Categorical univariate AFT models

Table 4.8 and 4.9 below shows the coefficients and the corresponding time ratio (TR), 95% confidence intervals and p-values for categorical univariate Weibull, Exponential, Log-logistic and Gamma AFT models, respectively. TR greater than 1, shows the prolonged survival time and TR less than 1 is related to a decrease in survival time (Swain and Grover, 2016).

Therefore, from Table 4.8, the univariate Weibull AFT model, revealed that gram (gram negative versus gram positive: TR=0.515, 95% CI=-1.221 to -0.104, p=0.0201), water (no water versus water: TR=0.621, 95% CI=-0.919 to -0.034, p=0.0350), electricity (no electricity versus electricity: TR=0.249, 95% CI=-2.271 to -0.511, p=0.0020), dwelling (rural versus urban: TR=0.319, 95% CI=-2.049 to -0.234, p=0.0137), employment (unemployed versus employed: TR=0.442, 95% CI=-1.479

to -0.155, p=0.0155) were significantly associated with the development of time-to-first peritonitis at 5% significant level. However, the model found that sex, race, anuric at baseline, dialysate, diastolic dysfunction, LVH, PHT, season, causes of ESRD, house type and education were statistically insignificant.

Thus, with respect to TR, patients with gram negative organisms during PD, have lesser chance of survival compared to their gram positive counterpart. We can also observe that patients with no water during PD, have lesser chance of survival as compared to patients with water. Similarly, we found that patients without electricity, have lesser chance of survival as compared to patients with electricity during PD. Moreover, patients coming from rural areas, were also found to have lesser survival chance as compared to patients coming from urban areas. The analysis further reveals that patients who are unemployed, have lesser survival chance as compared to patients who are employed.

The univariate exponential AFT model in Table 4.8, shows that water (no water versus water: TR=0.626, 95% CI=-0.892 to -0.045, p=0.0303), electricity (no electricity versus electricity: TR=0.246, 95% CI=-2.305 to -0.499, p=0.0024), dwelling (rural versus urban: TR=0.297, 95% CI=-2.212 to -0.206, p=0.0182), employment(unemployed versus employed: TR=0.437, 95% CI=-1.490 to -0.165, p=0.0143) are associated risk factors of time-to-first peritonitis episode. However, the model found that gram, sex, race, anuric at baseline, dialysate, diastolic dysfunction, LVH, PHT, season, causes of ESRD, house type and education were statistically insignificant. Thus, the interpretation of TR for significant factors in this model, can be done in the same fashion as those in the univariate Weibull AFT model.

Table 4.8: Univariate categorical Weibull and Exponential AFT models:

Covariate		Weibull AFT model					Exp AFT model				
	category	β	TR	LCL	UCL	P-VALUE	β	TR	LCL	UCL	P-VALUE
Gram	Gram-negative	-0.663	0.515	-1.221	-0.104	0.0201	-0.647	0.524	-1.324	0.030	0.0612
	Gram-positive (ref)	0.000					0.000				
Anuric at baseline	Anuric	0.246	1.279	-0.454	0.946	0.4909	0.244	1.276	-0.447	0.934	0.4894
	Non-Anuric (ref)	0.000					0.000				
Dialysate	Adcock-Icod	-0.257	0.773	-1.263	0.749	0.6171	-0.257	0.773	-1.261	0.747	0.6158
	Bicavera	-0.059	0.943	-0.693	0.575	0.8552	-0.059	0.493	-0.692	0.573	0.8546
	Dianeal-PD-Sol	0.009	1.009	-0.570	0.589	0.9734	0.009	1.009	-0.565	0.583	0.9752
	Stay-safe (ref)	0.000					0.000				
Dystolic-dysfn	No	-0.281	0.755	-0.743	0.182	0.2342	-0.292	0.747	-0.799	0.216	0.2606
	Yes (ref)	0.000					0.000				
LVH	No	-0.286	0.751	-0.718	0.147	0.1956	-0.298	0.742	-0.773	0.178	0.2196
	Yes (ref)	0.000					0.000				
PHT	No	-0.529	0.589	-1.353	0.294	0.2074	-0.543	0.581	-1.453	0.368	0.2429
	Yes (ref)	0.000					0.000				
Season	Autumn	-0.139	0.870	-0.731	0.455	0.6488	-0.138	0.871	-0.729	0.453	0.6481
	Summer	-0.228	0.796	-0.823	0.367	0.4524	-0.227	0.797	-0.818	0.364	0.4513
	Winter	-0.439	0.645	-1.037	0.159	0.1507	-0.437	0.646	-1.028	0.154	0.1473
	Spring (ref)	0.000					0.000				
Sex	Female	-0.133	0.875	-0.557	0.291	0.5389	-0.131	0.877	-0.549	0.287	0.5387
	Male (ref)	0.000					0.000				
Race	Black	-1.131	0.323	-2.524	0.261	0.1114	-1.143	0.319	-2.545	0.259	0.1101
	Indian	-2.315	0.099	-4.688	0.058	0.0558	-2.324	0.098	-4.724	0.077	0.0578
	White (ref)	0.000					0.000				
Water	No	-0.477	0.621	-0.919	-0.034	0.0350	-0.468	0.626	-0.892	-0.045	0.0303
	Yes (ref)	0.000				0.000					
Electricity	No	-1.391	0.249	-2.271	-0.511	0.0020	-1.402	0.246	-2.305	-0.499	0.0024
	Yes (ref)	0.000					0.000				
Employment	Scholar	-0.806	0.447	-1.756	0.145	0.0968	-0.808	0.446	-1.774	0.158	0.1012
	Unemployed	-0.817	0.442	-1.479	-0.155	0.0155	-0.828	0.437	-1.490	-0.165	0.0143
	Employed (ref)	0.000					0.000				
Dwelling	Rural	-1.141	0.319	-2.049	-0.234	0.0137	-1.209	0.298	-2.212	-0.206	0.0182
	Urban (ref)	0.000					0.000				
House type	Formal	0.608	1.837	-0.537	1.752	0.2982	0.607	1.835	-0.545	1.759	0.3015
	Informal (ref)	0.000					0.000				
Education	No formal education	-0.813	0.444	-2.117	0.491	0.2218	-0.814	0.443	-2.129	0.500	0.2247
	Primary	-0.005	0.995	-0.946	0.936	0.9922	-0.005	0.995	-0.954	0.944	0.9919
	Secondary	-0.069	0.933	-0.981	0.843	0.8819	-0.071	0.931	-0.990	0.848	0.8797
	Tertiary (ref)	0.000					0.000				
Causes of ESRD	Diabetes	0.133	1.142	-0.759	1.025	0.7705	0.129	1.138	-0.752	1.012	0.7727
	Glomerolo	0.410	1.507	-0.655	1.476	0.4506	0.403	1.496	-0.645	1.450	0.4511
	others	-0.012	0.988	-0.742	0.718	0.9739	-0.013	0.987	-0.735	0.709	0.9715
	Unknown	0.085	1.089	-0.413	0.582	0.7379	0.083	1.087	-0.408	0.574	0.7412
	hypertension (ref)	0.000					0.000				

Abbreviations: Ref: Reference, LVH: Left ventricular hypertrophy, PHT: Pulmonary Hypertension, Dysfn: Dysfunction, Season: Season of the catheter insertion, PD-Sol: Peritoneal dialysis solution

The univariate log-logistic AFT model as shown in Table 4.9, revealed that season (winter versus spring: TR=0.440, 95% CI=-1.436 to -0.205, p=0.0090), water (no water versus water: TR=0.614, 95% CI=-0.960 to -0.016, p=0.0427), electricity (no electricity versus electricity: TR=0.0129, 95% CI=-2.193 to -0.260, p=0.0129), Dwelling (rural versus urban: TR=0.319, 95% CI=-2.049 to -0.234, P=0.0137), employment (unemployed versus employed: TR=0.487, 95% CI=-1.375 to -0.065, p=0.0313) are significantly associated with development of time-to-first peritonitis at 5% significance level. However, the model found that gram, sex, race, anuric at baseline, dialysate, diastolic dysfunction, LVH,PHT, causes of ESRD, house type

and education were statistically insignificant.

These findings indicates that patients who insert their catheter during winter, have lesser chance of survival as compared to patients with catheter insertion during spring season, with summer and autumn showing no significance association with time-to-first peritonitis. The interpretation of TR for other significant factors in this model can be done in the same fashion as those in the univariate Weibull AFT model.

The univariate gamma AFT model as shown in Table 4.9, revealed that season (winter versus spring: TR=0.425, 95% CI=-1.422 to -0.290, p=0.0030), water (no water versus water: TR=0.616, 95% CI=-0.954 to -0.014, p=0.0435), dwelling (rural versus urban: TR=0.341, 95% CI=-1.851 to -0.300, p=0.0065) are significantly associated with development of time-to-first peritonitis at 5% significance level. However, the model also found that gram, electricity, employment, sex, race, anuric at baseline, dialysate, diastolic dysfunction, LVH,PHT, causes of ESRD, house type and education were statistically insignificant.

Thus, patients who insert their catheter during winter, have lesser chance of survival as compared to patients with catheter insertion during spring season, with summer and autumn showing no significant association with time-to-first peritonitis. The interpretation of TR for other significant factors in this model can be done in the same fashion as those in the univariate Weibull AFT model.

Table 4.9: Univariate categorical Log-logistic and Gamma AFT models:

Covariate		Log-logistic AFT Model					Gamma AFT Model				
	category	β	**TR**	**LCL**	**UCL**	**P-VALUE**	β	**TR**	**LCL**	**UCL**	**P-VALUE**
Gram	Gram-negative	-0.542	0.582	-1.206	0.122	0.1094	-0.606	0.546	-1.228	0.017	0.0567
	Gram-positive (ref)	0.000					0.000				
Anuric at baseline	Anuric	0.359	1.432	-0.317	1.034	0.2978	0.469	1.598	-0.183	1.120	0.1587
	Non-Anuric (ref)	0.000					0.000				
Dialysate	Adcock-Icod	0.024	1.024	-1.055	1.103	0.9644	0.169	1.184	-0.996	1.334	0.7761
	Bicavera	-0.016	0.984	-0.725	0.692	0.9644	-0.125	0.882	-0.841	0.592	0.7330
	Dianeal-PD-Sol	0.158	1.171	-0.474	0.790	0.6233	0.076	1.079	-0.558	0.709	0.8147
	Stay-safe (ref)	0.000					0.000				
Dystolic-dysfn	No	-0.269	0.764	-0.763	0.223	0.2830	-0.287	0.751	-0.768	0.194	0.2427
	Yes (ref)	0.000					0.000				
LVH	No	-0.303	0.739	-0.771	0.165	0.2041	-0.384	0.681	-0.843	0.076	0.1019
	Yes (ref)	0.000					0.000				
PHT	No	-0.378	0.685	-1.285	0.529	0.4139	-0.185	0.831	-1.029	0.659	0.6681
	Yes (ref)	0.000					0.000				
Season	Autumn	-0.109	0.897	-0.699	0.480	0.7159	0.011	1.011	-0.554	0.576	0.9703
	Summer	-0.305	0.737	-0.918	0.308	0.3299	-0.399	0.671	-0.978	0.179	0.1760
	Winter	-0.820	0.440	-1.436	-0.205	0.0090	-0.856	0.425	-1.422	-0.290	0.0030
	Spring (ref)	0.000					0.000				
Sex	Female	-0.235	0.791	-0.686	0.215	0.3059	-0.274	0.760	-0.715	0.167	0.227
	Male (ref)	0.000					0.000				
Race	Black	-1.035	0.355	-2.337	0.267	0.1191	-0.728	0.483	-1.842	0.386	0.2003
	Indian	-1.788	0.167	-4.113	0.538	0.1320	-1.252	0.286	-3.867	1.363	0.3480
	White (ref)	0.000					0.000				
Water	No	-0.488	0.614	-0.960	-0.016	0.0427	-0.484	0.616	-0.954	-0.014	0.0435
	Yes (ref)	0.000					0.000				
Electricity	No	-1.227	0.293	-2.193	-0.260	0.0129	-1.039	0.354	-2.191	0.114	0.0773
	Yes (ref)	0.000					0.000				
Employment	Scholar	-0.606	0.546	-1.533	0.321	0.2001	-0.454	0.635	-1.378	0.469	0.3353
	Unemployed	-0.719	0.487	-1.375	-0.065	0.0313	-0.583	0.558	-1.220	0.053	0.0725
	Employed (ref)	0.000					0.000				
Dwelling	Rural	-1.141	0.319	-2.049	-0.234	0.0137	-1.076	0.341	-1.851	-0.300	0.0065
	Urban (ref)	0.000					0.000				
House type	Formal	0.689	1.992	-0.512	1.889	0.2611	0.539	1.714	-0.674	1.751	0.3839
	Informal (ref)	0.000					0.000				
Education	No formal education	-0.632	0.532	-2.030	0.766	0.3754	-0.352	0.703	-1.749	1.045	0.6212
	Primary	-0.018	0.982	-1.007	0.969	0.9709	0.133	1.142	-0.789	1.056	0.7772
	Secondary	0.144	1.155	-0.817	1.105	0.7691	0.298	1.347	-0.616	1.211	0.5228
	Tertiary (ref)	0.000					0.000				
Causes of ESRD	Diabetes	0.196	1.217	-0.759	1.150	0.6876	0.121	1.129	-0.819	1.063	0.8004
	Glomerolo	0.274	1.315	-0.780	1.327	0.6108	0.242	1.274	-0.714	1.198	0.6201
	others	-0.123	0.884	-0.195	0.669	0.7617	-0.102	0.903	-0.887	0.683	0.7990
	Unknown	0.038	1.039	-0.502	0.577	0.8907	-0.081	0.922	-0.634	0.472	0.7750
	hypertension (ref)	0.000					0.000				

Abbreviations: Ref: Reference, LVH: Left ventricular hypertrophy, PHT: Pulmonary Hypertension, Dysfn: Dysfunction, Season: Season of the catheter insertion, PD-Sol: Peritoneal dialysis solution

4.4.2 Summary of the univariate categorical AFT models: Additional potential factors

As stated previously, the univariate AFT analysis was aimed to identify additional significant factors which contributes to development of first peritonitis episode. Therefore, the univariate categorical Weibull AFT model, found that gram, water, electricity, employment and dwelling were significant. The univariate categorical exponential AFT model, found that water, electricity, employment and dwelling were significant. The univariate categorical log-logistic AFT model, found that season, water, electricity, employment and dwelling were significant and lastly, season, wa-

ter and dwelling were found to be significant in the univariate categorical gamma AFT model. Therefore, based on these findings, water, electricity, dwelling and season will be entered in the multivariate AFT models. Thus, we can conclude that no additional factors was found, since all these identified factors were already identified by Cox PH model. However, as stated before, due to high multicollinearity, factor gram and employment, which were again found to be significant using some of the univariate AFT models, will not be included in the multivariate AFT models.

4.4.3 Continuous univariate AFT models

Table 4.10 and 4.11 below shows the coefficients and the corresponding time ratio (TR), 95% confidence intervals and p-values for continuous univariate Weibull, Exponential, Log-logistic and Gamma AFT models, respectively.

Therefore, the univariate Weibull AFT model as shown in Table 4.10, revealed that ferritin (TR=0.999, 95% CI=-0.001 to -0.000, p=0.0476), Hb (TR=1.147, 95% CI=0.036 to 0.238, p=0.0079), duration on PD (TR=1.023, 95% CI=0.011 to 0.035, p=0.0002) and distance to dialysis center (TR=0.997, 95% CI=-0.006 to -0.000, p=0.0419) are significantly associated with development of time-to-first peritonitis at 5% significance level. However, the model found that age at baseline, weight, Alk-Phosp, BMI, cholestorol, CaCorr, DBP, No anti-HTN drugs, PTH, Pi, SBP. Transat, eGFR MDRD, Albumin, number of rooms and number of people were statistically insignificant at 5% significance level.

Thus, based on TR, we can state that a unit increase in ferrintin of a PD patient, correspond to a decrease in survival time. Similarly, we observe that a unit increase in Hb and duration on PD of a PD patient, respectively, would improve or prolong the estimated survival time. TR further reveals that a unit increase in distance to dialysis center of a PD patient, correspond to a decrease in survival time.

Furthermore, the univariate exponential AFT model as shown in Table 4.10, re-

vealed that Hb (TR=1.148, 95% CI=0.037 to 0.239, p=0.0076), duration on PD (TR=1.021, 95% CI=0.008 to 0.034, p=0.0016) and distance to dialysis center (TR=0.997, 95% CI=-0.006 to -0.000, p=0.0451) are significantly associated with the development of time-to-first peritonitis at 5% significance level. However, the model found that age at baseline, weight, Alk-Phosp, BMI, cholestorol, CaCorr, DBP, ferritin, No anti-HTN drugs, PTH, Pi, SBP. Tran-sat, eGFR MDRD, Albumin, number of rooms and number of people were statistically insignificant at 5% significance level. Thus, the interpretation of TR for significant factors in this model, can be done in the same fashion as those in the univariate Weibull AFT model.

Table 4.10: Univariate continuous Weibull and Exponential AFT models:

Covariate		**Weibull**	**AFT**	**model**			**Exp**	**AFT**	**model**	
	β	**TR**	**LCL**	**UCL**	**P-VALUE**	β	**TR**	**LCL**	**UCL**	**P-VALUE**
Age at baseline	0.005	1.005	-0.014	0.024	0.6039	0.005	1.005	-0.014	0.024	0.605
Weight	0.002	1.002	-0.014	0.018	0.8035	0.002	1.002	-0.011	0.014	0.7937
Alk-Phosp ($mmol/L$)	0.002	1.002	-0.002	0.006	0.2968	0.002	1.002	-0.002	0.005	0.2913
BMI (kg/m^2)	0.022	1.022	-0.022	0.067	0.3192	0.022	1.022	-0.022	0.067	0.3232
Cholestorol ($mmol/L$)	-0.003	0.997	-0.231	0.225	0.9805	-0.002	0.998	-0.037	0.233	0.9861
CaCorr ($mmol/L$)	0.135	1.145	-0.632	0.902	0.7304	0.135	1.145	-0.635	0.902	0.7310
DBP ($mmhg$)	-0.004	0.996	-0.016	0.008	0.5251	-0.004	0.996	-0.015	0.008	0.5057
Ferritin(ug/l)	-0.001	0.999	-0.001	-0.000	0.0476	-0.001	0.999	-0.001	0.000	0.0520
Hb (g/dL)	0.137	1.147	0.036	0.238	0.0079	0.138	1.148	0.037	0.239	0.0076
No anti-HTN drugs	-0.108	0.898	-0.296	0.081	0.2626	-0.107	0.899	-0.302	0.089	0.2845
PTH (ng/L)	0.000	1.000	-0.000	0.001	0.3219	0.000	1.000	-0.000	0.001	0.3254
Pi ($mmol/L$)	-0.014	0.986	-0.264	0.236	0.9102	-0.014	0.986	-0.268	0.239	0.9115
SBP ($mmHg$)	-0.007	0.993	-0.015	0.002	0.1082	-0.007	0.993	-0.015	0.001	0.1041
Trans-Sat (%)	-0.008	0.992	-0.019	0.002	0.1298	-0.009	0.991	-0.020	0.003	0.1382
eGFR MDRD ($mL/min/1.73m^2$)	0.012	1.012	-0.047	0.071	0.6925	0.012	1.012	-0.047	0.071	0.6926
Duration on PD	0.023	1.023	0.011	0.035	0.0002	0.021	1.021	0.008	0.034	0.0016
Albumin	0.019	1.019	-0.015	0.053	0.2836	0.019	1.019	-0.016	0.054	0.2903
Distance to dialysis center (km)	-0.003	0.997	-0.006	-0.000	0.0419	-0.003	0.997	-0.006	-0.000	0.0451
Number of rooms	-0.009	0.991	-0.117	0.099	0.8709	-0.009	0.991	-0.119	0.101	0.8727
Number of people	-0.034	0.967	-0.119	0.053	0.4471	-0.034	0.967	-0.119	0.053	0.4471

Abbreviations: Alk Phos: Alkaline phosphatase, BMI: Body mass index, CaCorr: calcium corrected, DBP: Diastolic blood pressure, Hb: hemoglobin, Anti-HTN: anti hypertensive, PTH: Parathyroid, hormone, Pi: phosphorus, SBP: systolic blood pressure, Trans-Sat: saturation of transferrin, eGFR MDRD: glomerular filtration rate according to MDRD formula, PD: Peritoneal dialysis

The univariate log-logistic AFT model as shown in Table 4.11, revealed that ferritin (TR=0.999, 95% CI=-0.001 to -0.000, p=0.0458), Hb (TR=1.206, 95% CI=0.084 to 0.290, p=0.0004) and duration on PD (TR=1.021, 95% CI=0.009 to 0.032, p=0.0003) are significantly associated with the development of time-to-first peritonitis at 5% significance level. However, the model found that age at baseline, weight, Alk-Phosp, BMI, cholestorol, CaCorr, DBP, No anti-HTN drugs, PTH, Pi, SBP. Tran-sat, eGFR MDRD, Albumin, distance to dialysis center, number of rooms and number of people were statistically insignificant at 5% significance level. Thus, the inter-

pretation of TR for significant factors in this model can be done in the same fashion as those in the univariate Weibull AFT model.

The univariate gamma AFT model as shown in Table 4.11, revealed that Hb (TR=1.241, 95% CI=0.116 to 0.316, p=0.0001) and duration on PD (TR=1.019, 95% CI=0.009 to 0.030, p=0.0002) are significantly associated with the development of time-to-first peritonitis at 5% significance level. However, the model found that age at baseline, weight, Alk-Phosp, BMI, cholestorol, CaCorr, DBP, ferritin, No anti-HTN drugs, PTH, Pi, SBP. Tran-sat, eGFR MDRD, Albumin, distance to dialysis center, number of rooms and number of people were statistically insignificant at 5% significance level. Thus, the interpretation of TR for significant factors in this model can be done in the same fashion as those in the univariate Weibull AFT model.

Table 4.11: Univariate continuous Log-logistic and Gamma AFT models:

Covariate	**Log-logistic AFT model**					**Gamma AFT model**				
	β	**TR**	**LCL**	**UCL**	**P-VALUE**	β	**TR**	**LCL**	**UCL**	**P-VALUE**
Age at baseline	0.002	1.002	-0.017	0.021	0.8528	-0.003	0.997	-0.022	0.016	0.7409
Weight	0.003	1.003	-0.014	0.021	0.6935	0.002	1.002	-0.014	0.019	0.7805
Alk-Phosp ($mmol/L$)	0.002	1.002	-0.002	0.005	0.2845	0.001	1.001	-0.002	0.004	0.4278
BMI (kg/m^2)	0.014	1.014	-0.032	0.059	0.5640	0.012	1.012	-0.034	0.057	0.6080
Cholestorol ($mmol/L$)	0.054	1.055	-0.183	0.291	0.6528	0.005	1.005	-0.191	0.201	0.9626
CaCorr ($mmol/L$)	0.109	1.115	-0.627	0.846	0.7716	0.228	1.256	-0.496	0.953	0.5371
DBP ($mmhg$)	0.001	1.001	-0.012	0.014	0.8734	0.003	1.003	-0.009	0.016	0.6197
Ferritin (ug/l)	-0.001	0.999	-0.001	-0.000	0.0458	-0.001	0.999	-0.001	-0.000	0.0663
Hb (g/dL)	0.187	1.206	0.084	0.290	0.0004	0.216	1.241	0.116	0.316	0.0001
No anti-HTN drugs	-0.004	0.996	-0.196	0.188	0.9686	0.035	1.036	-0.177	0.247	0.743
PTH (ng/L)	0.000	1.000	-0.000	0.001	0.2226	0.001	1.001	-0.000	0.001	0.1317
Pi ($mmol/L$)	-0.019	0.981	-0.278	0.239	0.8813	-0.028	0.972	-0.257	0.201	0.8112
SBP ($mmHg$)	-0.004	0.996	-0.012	0.005	0.4195	-0.001	0.999	-0.009	0.008	0.8507
Trans-Sat (%)	-0.009	0.991	-0.022	0.003	0.1358	-0.008	0.992	-0.021	0.004	0.1935
eGFR MDRD ($mL/min/1.73m^2$)	0.010	1.010	-0.059	0.080	0.7700	0.021	1.021	-0.053	0.094	0.5790
Duration on PD	0.021	1.021	0.009	0.032	0.0003	0.019	1.019	0.009	0.030	0.0002
Albumin	0.024	1.024	-0.012	0.060	0.1950	0.025	1.025	-0.011	0.061	0.1670
Distance to dialysis center (km)	-0.003	0.997	-0.006	0.001	0.1228	-0.002	0.998	-0.005	0.001	0.1532
Number of rooms	0.019	1.019	-0.081	0.120	0.6995	0.018	1.018	-0.079	0.115	0.7120
Number of people	-0.007	0.993	-0.101	0.087	0.8835	-0.007	0.993	-0.105	0.092	0.8962

Abbreviations: Alk Phos: Alkaline phosphatase, BMI: Body mass index, CaCorr: calcium corrected, DBP: Diastolic blood pressure, Hb: hemoglobin, Anti-HTN: anti hypertensive,PTH: Parathyroid hormone, Pi: phosphorus, SBP: systolic blood pressure, Trans-Sat: saturation of transferrin, eGFR MDRD: glomerular filtration rate according to MDRD formula, PD: Peritoneal dialysis

4.4.4 Summary of the univariate continuous AFT models: Additional potential factors

The univariate continuous Weibull AFT model, found that ferritin, Hb, duration on PD and distance to dialysis center were significant. The univariate continuous

exponential AFT model, found that Hb, duration on PD and distance to dialysis center were significant. The univariate continuous log-logistic AFT model, found that ferritin, Hb and duration on PD were significant and lastly, Hb and duration on PD were found to be significant in the univariate continuous gamma AFT model. Therefore, based on these findings, ferritin, Hb, duration on PD and distance to dialysis center will be entered in the multivariate AFT models. Hence, we can conclude that ferritin, duration on PD and distance to dialysis center are additional factors.

4.4.5 Multivariate Weibull, Exponential, Log-logistic and Gamma AFT Models

This section fits full multivariate AFT models, by presenting the multivariate effect of risk factors on time-to-first peritonitis episode. The AFT models include all potential risk factors which were identified by Cox PH model and univariate AFT models. As stated previously, water, electricity, dwelling, season, Hb, ferritin, duration on PD and distance to dialysis center will be entered in the multivariate AFT models. Tran-sat was the only factor, which was entered in the final multivariate Cox PH model, but were not found to be significant by any of the univariate continuous AFT models. However, it will be entered in the multivariate AFT models. Table 4.12 and 4.13, shows the results of multivariate Weibull, exponential, log-logistic and gamma AFT Models. This results consists of TR, p-values and 95% confidence intervals for these factors.

The multivariate Weibull AFT model as shown in Table 4.12, shows that water (no water versus water: TR=0.551, 95% CI=-0.997 to -0.195, p=0.0036), electricity (no electricity versus electricity: TR=0.368, 95% CI=-1.823 to -0.179, p=0.0170), dwelling (rural versus urban: TR=0.184, 95% CI=-3.166 to -0.217, p=0.0246), Hb (TR=1.135, 95% CI=0.036 to 0.218, p=0.0061) and duration on PD (TR=1.017, 95% CI=0.006 to 0.028, p=0.0019) are significantly associated with time-to-first

peritonitis. However, season, tran-sat, distance to dialysis center and ferritin were statistically insignificant.

With respect to the observed TR, the model indicates that PD patients with no water, appear not to prolong time-to-first peritonitis or correspond to a decrease in survival time during PD. Therefore, suggesting that patients with no water, have higher chance of developing first peritonitis episode than patients with water. Similarly, the model shows that PD patients with no electricity, appear not to prolong time-to-first peritonitis or correspond to a decrease in survival time during PD. Therefore, suggesting that patients without electricity, have a higher chance of developing first peritonitis episode than patients with electricity. The model also shows that PD patients coming from rural areas, appear to have a shorter survival chance than PD patients who are from urban areas. These findings suggest that rural dwellers have a higher chance of developing first peritonitis episode as compared to urban dwellers.

TR also reveals that a unit increase in Hb of a PD patient, will increase or prolong the survival time of a PD patient. This results suggest that patients with higher Hb will have a better survival or a lower chance of developing first peritonitis episode than those with lower Hb. Furthermore, the TR for duration on PD indicates that a unit increase in PD duration of a patient, would improve the estimated survival time. Therefore, suggesting that patients with shorter PD duration have a lesser chance of survival or higher chance of developing peritonitis for the first time.

The multivariate exponential AFT model as shown in Table 4.12, revealed that water (no water versus water: TR=0.551, 95% CI=-0.997 to -0.195, p=0.0036), electricity (no electricity versus electricity: TR=0.368, 95% CI=-1.823 to -0.179, p=0.0170), dwelling (rural versus urban: TR=0.184, 95% CI=-3.166 to -0.217, p=0.0246), Hb (TR=1.135, 95% CI=0.036 to 0.218, p=0.0061) and duration on PD (TR=1.017, 95% CI=0.006 to 0.028, p=0.0019) are significantly associated with time-to-first peritonitis. But other characteristics such as season, tran-sat, distance to dialysis

center and ferritin were statistically insignificant. This findings are the same as the ones obtained by multivariate Weibull AFT model. Furthermore, slightly differences with respect to TR of water, electricity, dwelling, Hb and duration on PD were observed as compared to Weibull model. However, similar conclusions about TR was made.

Table 4.12: Multivariate Weibull and Exponential AFT models:

Covariate		Weibull AFT Model					Exp AFT Model				
	category	β	TR	LCL	UCL	P-VALUE	β	TR	LCL	UCL	P-VALUE
intercept		3.871		2.067	5.675	0.0001	4.378		1.926	6.831	0.0005
Water	No	-0.596	0.551	-0.997	-0.195	0.0036	-0.708	0.493	-1.260	-0.156	0.0119
	Yes (ref)	0.000					0.000				
Electricity	No	-1.001	0.368	-1.823	-0.179	0.0170	-1.152	0.316	-2.245	-0.059	0.0388
	Yes (ref)	0.000					0.000				
Dwelling	Rural	-1.691	0.184	-3.166	-0.217	0.0246	-2.216	0.109	-4.231	-0.201	0.0311
	Urban (ref)	0.000					0.000				
Season	Autumn	-0.156	0.856	-0.688	0.376	0.5654	-0.194	0.824	-0.195	0.527	0.5980
	Summer	-0.225	0.799	-0.827	0.377	0.4640	-0.245	0.783	-1.075	0.586	0.5634
	Winter	-0.524	0.592	-1.082	0.035	0.0661	-0.555	0.574	-1.322	0.213	0.1567
	Spring (ref)	0.000					0.000				
Hb (g/dL)		0.127	1.135	0.036	0.218	0.0061	0.155	1.168	0.031	0.278	0.0140
Trans-Sat (%)		-0.007	0.993	-0.017	0.003	0.1520	-0.008	0.992	-0.021	0.005	0.2444
Duration on PD		0.017	1.017	0.006	0.028	0.0019	0.015	1.015	0.000	0.029	0.0462
Distance to dialysis center (km)		0.000	1.000	-0.003	0.003	0.9863	0.000	1.000	-0.004	0.004	0.9904
Ferritin (ug/l)		-0.000	1.000	-0.001	0.000	0.1966	0.000	1.000	-0.001	0.000	0.3646
Scale		0.717		0.581	0.885		1.000		1.000	1.000	
Weibull shape		1.394		1.129	1.720		1.000		1.000	1.000	

Abbreviations: Ref: reference, Season: season of the catheter insertion, Hb: hemoglobin, Trans-Sat: saturation of transferrin

Under the multivariate log-logistic AFT model shown in Table 4.13, it was found that water (no water versus water: TR=0.589, 95% CI=-0.964 to -0.092, p=0.0177), electricity (no electricity versus electricity: TR=0.364, 95% CI=-1.847 to -0.175, p=0.0177), dwelling (rural versus urban: TR=0.207, 95% CI=-2.739 to -0.406, p=0.0083), Hb (TR=1.166, 95% CI=0.056 to 0.251, p=0.0021) and duration on PD (TR=1.017, 95% CI=0.007 to 0.028, p=0.0017) are significantly associated with time-to-first peritonitis. However, season, tran-sat, distance to dialysis center and ferritin were statistically insignificant. This findings are the same with the ones obtained by multivariate Weibull and exponential AFT models. Therefore, same conclusions about TR was made.

That is, PD patients with no water, appear not to prolong time-to-first peritonitis or correspond to a decrease in survival time during PD. Therefore, suggesting that patients with no water, have higher chance of developing first peritonitis episode than patients with water. Similarly, the model shows that PD patients with no elec-

tricity, appear not to prolong time-to-first peritonitis or correspond to a decrease in survival time during PD. Therefore, suggesting that patients without electricity, have higher chance of developing first peritonitis episode. The model also shows that PD patients coming from rural areas appear to have shorter survival chances than PD patients who are from urban areas. This findings suggest that rural dwellers have a higher chance of developing first peritonitis episode as compared to urban dwellers.

TR also reveals that a unit increase in Hb of a PD patient, will increase or prolong the survival time of a PD patient. This results suggest that patients with higher Hb will have a better survival or a lower chance of developing first peritonitis episode than patients with lower Hb. Furthermore, the TR for duration on PD indicates that a unit increase in PD duration of a patient, would improve the estimated survival time. Therefore, suggesting that patients with shorter PD duration have lesser chance of survival or higher chance of developing peritonitis for the first time.

The multivariate gamma model as shown in Table 4.13, reveals that dwelling (rural versus urban: TR=0.244, 95% CI=-2.234 to -0.588, p=0.0008), Hb (TR=1.190, 95% CI=0.068 to 0.281, p=0.0013) and duration on PD (TR=1.017, 95% CI=0.006 to 0.028, p=0.0026) are significantly associated with time-to-first peritonitis. However, water, electricity, season, tran-sat, distance to dialysis center and ferritin were statistically insignificant. The interpretation of TR for dwelling, Hb and duration on PD can be done in the same fashion as those in the multivariate Weibull and log-logistic AFT models.

Table 4.13: Multivariate Log-logistic and Gamma AFT models:

Covariate		Log-logistic AFT Model					Gamma AFT Model				
	category	β	TR	LCL	UCL	P-VALUE	β	TR	LCL	UCL	P-VALUE
intercept		3.078		1.527	4.629	0.0001	2.374		0.707	4.041	0.0052
Water	No	-0.528	0.589	-0.964	-0.092	0.0177	-0.462	0.630	-0.939	0.017	0.0586
	Yes (ref)	0.000					0.000				
Electricity	No	-1.011	0.364	-1.847	-0.175	0.0177	-0.855	0.425	-1.823	0.113	0.0835
	Yes (ref)	0.000					0.000				
Dwelling	Rural	-1.573	0.207	-2.739	-0.406	0.0083	-1.411	0.244	-2-234	-0.588	0.0008
	Urban (ref)	0.000					0.000				
Season	Autumn	-0.101	0.904	-0.618	0.416	0.7030	0.009	1.009	-0.536	0.554	0.9747
	Summer	-0.129	0.879	-0.742	0.483	0.6783	-0.053	0.948	-0.673	0.567	0.867
	Winter	-0.540	0.583	-1.123	0.043	0.0693	-0.471	0.624	-1.048	0.106	0.1097
	Spring (ref)	0.000					0.000				
Hb (g/dL)		0.154	1.166	0.056	0.251	0.0021	0.174	1.190	0.068	0.281	0.0013
Trans-Sat (%)		-0.006	0.994	-0.015	0.003	0.2065	-0.004	0.996	-0.014	0.007	0.4830
Duration on PD		0.017	1.017	0.007	0.028	0.0017	0.017	1.017	0.006	0.028	0.0026
Distance to dialysis center (km)		-0.001	0.999	-0.004	0.002	0.6742	-0.001	0.999	-0.004	0.002	0.5808
Ferritin (ug/l)		-0.000	1.000	-0.001	0.000	0.4743	-0.000	1.000	-0.001	0.000	0.7255
Scale		0.515		0.414	0.640		0.919		0.757	1.115	
shape							-0.537		-1.862	0.789	

Abbreviations: Ref: reference, Season: season of the catheter insertion, Hb: hemoglobin, Trans-Sat: saturation of transferrin

4.4.6 Comparison of AFT models using AIC-multivariate analysis

To test which parametric AFT model fits time-to-first peritonitis data set better, the AIC was used and any model with a smaller AIC fits the data set better. The results show that log-logistic appears to be appropriate AFT model according to AIC as compared to other three AFT models, since its AIC value of 223.466 is the least, as shown in Table 4.14. Although, we observe that its AIC value is slightly better than the AIC value of a gamma AFT model, which is 223.575. We also observe that Weibull and exponential AFT models are poor fits, since their AIC's of 228.019 and 234.306 are the largest. Hence, we can conclude that log-logistic AFT model is the best model in the multivariate analysis.

Table 4.14: AIC values for AFT models: Multivariate analysis

AFT MODEL	AIC VALUES
Weibull	228.091
Exponential	234.306
Log-logistic	223.466
Gamma	223.575

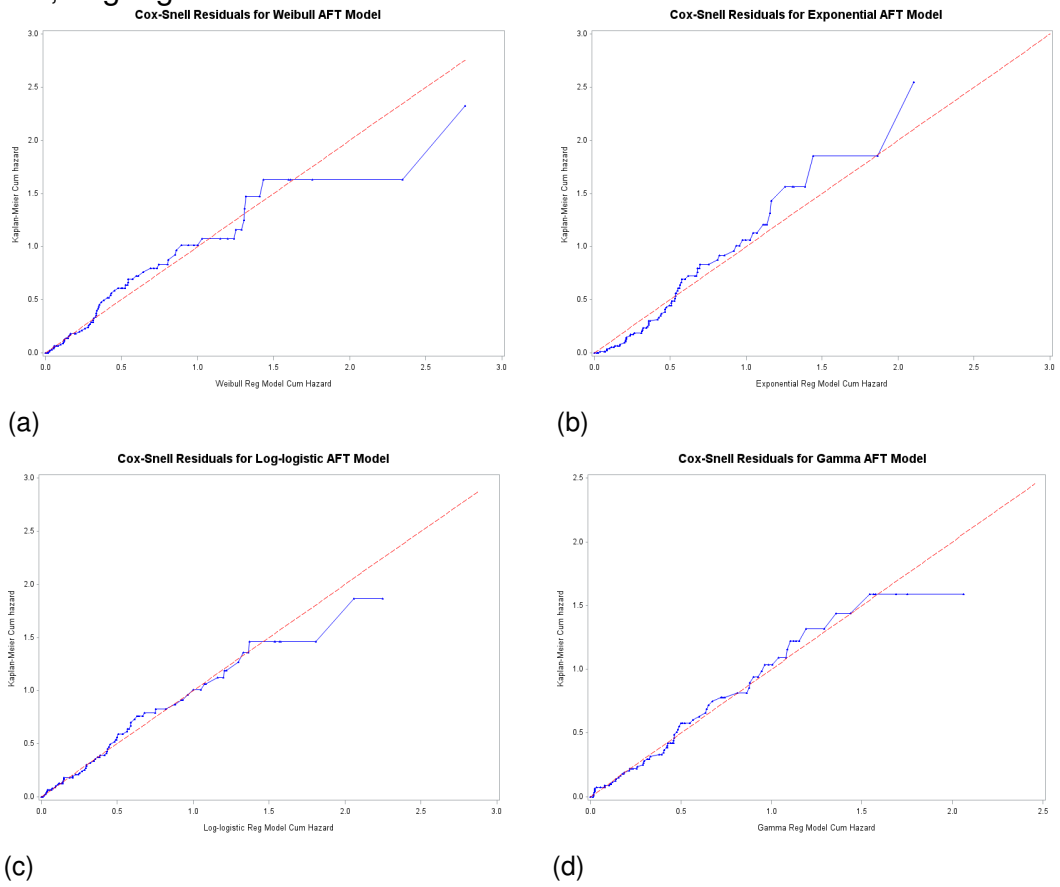
4.4.7 Test for the overall model fit: Cox-Snell residuals

The overall fit of Weibull, exponential, log-logistic and gamma AFT models are checked using the diagnostic plot of the Cox-Snell residuals. Using Cox-Snell residuals, if the plotted points lie on the line with an intercept zero and the slope unity, then it shows that the fitting of the model is good (Khanal et al., 2014; Jiezhi, 2009). In other words, if the fitted model is appropriate, the plot of $log(-logS(r_{CS_i}))$ vs $logr_{CS_i}$ should approximate a straight line with unit slope through the origin.

Therefore, the Cox-Snell residual plots in Figures 4.5 (a) and (b) for Weibull and exponential AFT models, show some evidence of systematic deviation from the straight line, and give some concern about the adequacy of these fitted models. Furthermore, from the Cox-Snell residual plots for both log-logistic and gamma AFT models in Figures 4.5 (c) and (d), we observe that for most part, the plotted points lie on the zero intercept and the unit slope.

Based on these diagnostic plots, there is no any reason to question or doubt the suitability of the fitted log-logistic and gamma AFT models to time-to-first peritonitis data set. However, although log-logistic and gamma AFT models provide reasonable fit, log-logistic AFT model appears to be appropriate AFT model as compared to other AFT models based on AIC. Hence, we conclude that based on AIC and Cox-Snell residual plots, the log-logistic AFT model is the best model fit as compared to other AFT models.

Figure 4.5: Cumulative hazard plot of the Cox-Snell residuals for Weibull, Exponential, Log-logistic and Gamma AFT models



4.5 Comparison of Cox PH model and log-logistic AFT model

This section will compare the selected multivariate AFT model with Cox PH model. Therefore, multivariate log-logistic AFT model was selected as a best AFT model based on AIC and Cox-Snell residuals. Hence, Table 4.15 compares multivariate log-logistic AFT model and multivariate Cox PH model based on AIC. As indicated previously, the model with the smallest AIC appears to be appropriate. The AIC value for the multivariate Cox PH model is very higher than the AIC value for log-logistic AFT model, indicating poor performance. Therefore, we can conclude that

log-logistic AFT model appears to be the best model for our dataset.

Table 4.15: AIC values for log-logistic AFT model and Cox PH model: Multivariate analysis

AFT MODEL	AIC VALUES
Log-logistic	223.466
Cox PH model	505.591

4.6 Discussion of the results

In the area of medical research, Cox PH model is the most widely used regression model for analysing survival data (Khanal et al., 2014). However, Jiezhi (2009) indicated that at most 5% of all studies utilising Cox PH model check PH assumption. If Cox PH assumption is checked, and do not hold, the model could lead to misleading conclusions. Alternatively, if that is the case, we can utilise the model where we stratify non-proportional factors. Furthermore, parametric AFT models such as Weibull, exponential, log-logistic and gamma are some common choices of analysis survival data even when the hazards are non-proportional.

In this study, we investigated time-to-first peritonitis data set using different survival analysis techniques. The aim was to apply and compare various survival analysis techniques and thereby identify the associated social, demographic and biological factors that contribute to first peritonitis episode in PD patients at Pietersburg Provincial Hospital. The categorical variables evaluated were gram, water, electricity, employment, dwelling, seasons, catheter removal, eGFR, sex, race, education, house type, anuric baseline, dialysate, diastolic dysfn, LVH, PHT and cause ESRD. The continuous variables were age baseline, alk phosp, BMI, cholesterol, CaCorr, DBP, distance KM, ferritin, Hb, number of rooms, number of people, PTH, Pi, Product, SBP, Trans Sat, eGFR2, eGFR MDRD, serum albumin. Cox PH model and Weibull, exponential, log-logistic and gamma AFT survival analysis techniques were used for analysis of this variables.

To build up multivariate Cox PH model, the univariate Cox PH model was used to test the association between each factor and time-to-first peritonitis. The threshold to enter each factor in the multivariate Cox PH model was p-value below 5%. Univariate Cox PH model found that gram, season, water, electricity, employment, dwelling, Hb and Tran-sat seemed to be involved in impacting the development of first peritonitis episode. However, due to high level of multicollinearity, gram organisms and employment were not included in the multivariate Cox PH model. Furthermore, before fitting multivariate Cox PH model, the study used survival curves (plot of log-log of survival against log of survival time) and supremum test to examine the PH assumption. The study found that PH assumption was not violated for any of the factors in the fitted Cox PH model.

Hence, we further explored the impact of season, water, electricity, dwelling, Hb and tran-sat towards the development of first peritonitis episode using multivariate Cox PH model. Multivariate analysis showed higher risk of developing first peritonitis episode for patients who had no water, no electricity, coming from rural areas and having lower Hb. However, the multivariate analysis also showed that season and tran-sat were not important factors towards the development of first peritonitis episode. After fitting the multivariate Cox PH model, deviance residual plots were used to assess the goodness of fit of the model. The PH model, based on deviance residual plots, seems to display a perfect fit. This indicated that there was no need for remedial measures such as stratifying.

To build up the multivariate AFT models, univariate analysis of all AFT models under consideration was applied to identify additional factors that could be associated with time-to-first peritonitis, which were not identified by Cox PH model. Interestingly, univariate AFT models found that ferritin, duration on PD and distance to dialysis center were additional factors. Hence, season, water, electricity, dwelling, Hb and tran-sat which were entered in the final Cox PH model as well as, ferritin, duration on PD and distance to dialysis center were all included in multivariate AFT models.

Furthermore, to select multivariate parametric AFT model that best fit time-to-first peritonitis data set, AIC and Cox-Snell residuals were used to examine the performance in the analysis. The results showed that the data strongly supported multivariate log-logistic AFT model, as the most suitable model than other AFT models based on AIC and Cox-snell residuals. The model was found to have the least AIC and showed that for most of the part, when Cox-Snell residuals was plotted, the plotted points lie on the zero intercept and the unit slope as compared to other AFT models. Therefore, there was no need to doubt the suitability of the fitted multivariate log-logistic AFT model.

Thus, under multivariate log-logistic AFT model, this study found that patients with no water and electricity, coming from rural areas, and having lower Hb with shorter duration on PD were significantly associated with the development of first peritonitis episode. Moreover, the model also revealed that season, tran-sat, ferritin and distance to dialysis center were not important factors towards the development of first peritonitis episode.

This study further compared multivariate Cox PH model and the selected log-logistic AFT model using AIC. It was found that log-logistic AFT model fit our data set better compared to Cox PH model, since the preferred model is the one with the smallest AIC value (Table 4.15). Hence, suggesting that an application of AFT models on time-to-first peritonitis data set at PKDC may be the best approach than Cox PH model. However, Stanley et al. (2016) emphasised that it must be appreciated that different methods may perform well under different settings, that is, the method that performed well in this study setting may prove inadequate in another setting.

Chapter 5

Conclusion and Recommendations

5.1 Introduction

This chapter present conclusions arising from the study, recommendations of what should be done to minimize the occurrence of first peritonitis episode, as well as the application of survival techniques in medical research. The strength and limitations of the study are also presented.

5.2 Conclusion

This prospective study was carried out in a single-dialysis center, from Limpopo province, South Africa, where the infection of first peritonitis episode is very high. In this study we have attempted to determine factors associated with the development of time-to-first peritonitis among kidney patients who were on PD using different survival analysis techniques. The study has used and validated both Cox PH model

and AFT models. We compared all AFT models using AIC statistical criteria and Cox-Snell residual plots. It was found that log-logistic AFT model looked to be a suitable AFT model. We further compared Cox PH model with the selected log-logistic AFT model using AIC. It was found that log-logistic AFT model provides a better description of time-to-first peritonitis in our study area. This study shows that AFT models have more convincing interpretation and offers more revealing results as compared to Cox PH model. Hence, suggesting that, although Cox PH assumption was not violated, utilising of Cox PH model may not be the finest approach.

In this study, the selected log-logistic AFT model revealed that availability of water, electricity, dwelling, haemoglobin status and duration on PD are significant risk factors towards the development of first peritonitis episode. Therefore, patients with no water and electricity, coming from rural background with low level of haemoglobin and shorter duration on PD are associated with high risk or hazard of developing peritonitis for the first time. Henceforth, it seems rational to theorize that PD patients presenting with a number of these factors, may be at higher risk of developing first peritonitis episode at PKDC.

5.3 Recommendations and areas for future research

In light of the study findings, the following recommendations are made. Firstly, it will be well for the investigators of health care field to consider AFT models instead of Cox PH model in further study related to time-to-first peritonitis. Secondly, we recommend and highlight that no matter how routine a method might be used in practice, it is a worthy practice to check goodness of fit of the model and where appropriate, the model should be compared with other models to find best fitting model. Thirdly, in some cases models are fitted by non-statisticians who either do not know how to evaluate model performance or might not even know that alter-

native methods exist. We recommend to such researchers that they must make an effort to ascertain their analyses with statisticians and to assist them to explore alternatives methods. Fourthly, because our study captured data on incidence PD patients from a single center, supplementary work with a complete data collection and multi-center involvement should be done to determine if the results of this study can be generalised to populations of other geographical areas or ethnicity. Lastly, future programs in this dialysis center and similar setting could consider taking a multidisciplinary approach towards developing home visits and regular investigation activities as part of their PD program. This is, to ensure that homes in rural areas have at least basic services such as water, electricity, and adequate hygiene.

5.4 Strengths of the study

To the best of our knowledge, this is the first study in South Africa, to utilise AFT models to fit time-to-first peritonitis data. The study was conducted at one of the biggest centers for PD in South Africa. As a result, the study has sufficient power to identify the associations and can also provide a risk prediction model firm parameter estimates. The study is prospectively designed, showing that there were no gaps in the amount and type of data that was collected, or no inherent problems of missing records. This made it possible to assess all relevant social and biological parameters known to be associated with peritonitis. Compared with other studies cohorts reported in literature, our study patients had a longer duration of follow-up period.

5.5 Limitation of the study

This study has a number of limitations. First and foremost, its status as a single-center study, which automatically indicates that although PKDC is one of the biggest center, the fact that is a single-center study means that small data was considered. Other limitations are that the data about peritoneal membrane characteristics were not available due to lack of proper tools to really evaluate the membrane. PD adequacy and RRF data were also not obtained which could have shed light on the effect of maintaining RRF. In addition, our study did not contain the data on exit-site catheter care, dedicated room to dialysis operation and patient's compliance, all the factors that may perhaps contribute to peritonitis. Regardless of this limitations, this population offers a reliable estimate of the effect of social, demographic and biological factors that contributes to first peritonitis episode on PD during a particular observation period.

REFERENCES

ABDU, A., LADEIRA, N., NAIDOO, S., AND NAICKER, S. (2011). The nutritional status of continuous ambulatory peritoneal dialysis patients at a johannesburg hospital. *South African Journal of Clinical Nutrition*, **24** (3), 150–153.

ABE, S., OBATA, Y., SATO, S., MUTA, K., KITAMURA, M., KAWASAKI, S., HIROSE, M., URAMATSU, T., AND NISHINO, T. (2016). Causative bacteria and risk factors for peritoneal dialysis-related peritonitis: A retrospective study. *Acta medica Nagasakiensia*, **60** (4), 145–151.

ANSIN, E. (2015). An evaluation of the cox-snell residuals. *Uppsala Universitet*.

BARONE, R., CÁMPORA, M., GIMENEZ, N., RAMIREZ, L., PANESE, S., AND SANTOPIETRO, M. (2011). Continuous ambulatory peritoneal dialysis versus automated peritoneal dialysis and peritonitis in the short and very long term at risk. *In Advances in peritoneal dialysis. Conference on Peritoneal Dialysis*, volume 28. pp. 44–49.

BÊLAVŜKOVÁ, S., FIVŜEROVÁ, E., AND KRUPIVŜKOVÁ, S. (2013). Study of bootstrap estimates in cox regression model with delayed entry. *Acta Universitatis Palackianae Olomucensis. Facultas Rerum Naturalium. Mathematica*, **52** (2), 21–30.

BENDER, F., BERNARDINI, J., AND PIRAINO, B. (2006). Prevention of infectious complications in peritoneal dialysis: best demonstrated practices. *Kidney International*, **70**, S44–S54.

BERNARDINI, J., BENDER, F., FLORIO, T., SLOAND, J., PALMMONTALBANO, L., FRIED, L., AND PIRAINO, B. (2005). Randomized, double-blind trial of antibiotic exit site cream for prevention of exit site infection in peritoneal dialysis patients. *Journal of the American Society of Nephrology*, **16** (2), 539–545.

BLAKE, P. G. (2001). Integrated end-stage renal disease care: the role of peritoneal dialysis. *Nephrology Dialysis Transplantation*, **16** (suppl_5), 61–66.

BOEHM, M., VÉCSEI, A., AUFRICHT, C., MUELLER, T., CSAICSICH, D., AND ARBEITER, K. (2005). Risk factors for peritonitis in pediatric peritoneal dialysis: a single-center study. *Pediatric Nephrology*, **20** (10), 1478–1483.

BOESCHOTEN, E. W., TER WEE, P. M., AND DIVINO, J. (2006). Peritoneal dialysis-related infections recommendations 2005—an important tool for quality improvement. *Nephrology Dialysis Transplantation*, **21** (suppl 2), ii31–ii33.

BORUCKA, J. (2014). Methods for handling tied events in the cox proportional hazard model. *Studia Oeconomica Posnaniensia*, **2** (2), 91–106.

BROWN, M. C. (2012). *Peritoneal dialysis in Scotland: an analysis of complications and outcomes in a contemporary national cohort*. Ph.D. thesis, University of Glasgow.

BUIS, M. L. (2006). An introduction to survival analysis. *Department of Social Research Methodology Vrije Universiteit Amsterdam [Online]*.

BUNKE, C. M., BRIER, M. E., AND GOLPER, T. A. (1997). Outcomes of single organism peritonitis in peritoneal dialysis: gram negatives versus gram positives in the network 9 peritonitis study. *Kidney international*, **52** (2), 524–529.

CHERN, Y.-B., HO, P.-S., KUO, L.-C., AND CHEN, J.-B. (2013). Lower education level is a major risk factor for peritonitis incidence in chronic peritoneal dialysis patients: a retrospective cohort study with 12-year follow-up. *Peritoneal dialysis international*, **33** (5), 552–558.

CHIDAMBARAM, M., BARGMAN, J. M., QUINN, R. R., AUSTIN, P. C., HUX, J. E., AND LAUPACIS, A. (2011). Patient and physician predictors of peritoneal dialysis technique failure: a population based, retrospective cohort study. *Peritoneal Dialysis International*, **31** (5), 565–573.

CHO, Y., BADVE, S. V., HAWLEY, C. M., MCDONALD, S. P., BROWN, F. G., BOUDVILLE, N., WIGGINS, K. J., BANNISTER, K. M., CLAYTON, P., AND JOHNSON, D. W. (2012). The effects of living distantly from peritoneal dialysis units on peritonitis risk, microbiology, treatment and outcomes: a multi-centre registry study. *BMC nephrology*, **13** (1), 41.

CHO, Y., BADVE, S. V., HAWLEY, C. M., MCDONALD, S. P., BROWN, F. G., BOUDVILLE, N., WIGGINS, K. J., BANNISTER, K. M., CLAYTON, P. A., AND JOHNSON, D. W. (2011). Seasonal variation in peritoneal dialysis-associated peritonitis: a multi-centre registry study. *Nephrology Dialysis Transplantation*, **27** (5), 2028–2036.

CHO, Y. AND JOHNSON, D. W. (2014). Peritoneal dialysis–related peritonitis: Towards improving evidence, practices, and outcomes. *American Journal of Kidney Diseases*, **64** (2), 278–289.

CHOW, K. M., SZETO, C. C., CHEUNG, K. K.-T., LEUNG, C. B., WONG, S. S.-H., LAW, M. C., HO, Y. W., AND LI, P. K.-T. (2006). Predictive value of dialysate cell counts in peritonitis complicating peritoneal dialysis. *Clinical Journal of the American Society of Nephrology*, **1** (4), 768–773.

CHOW, K. M., SZETO, C. C., LEUNG, C. B., LAW, M. C., AND LI, P. K.-T. (2005). Impact of social factors on patients on peritoneal dialysis. *Nephrology Dialysis Transplantation*, **20** (11), 2504–2510.

COX, D. R. AND SNELL, E. (1971). On test statistics calculated from residuals. *Biometrika*, **58** (3), 589–594.

CRABTREE, J. H., BURCHETTE, R. J., SIDDIQI, R. A., HUEN, I. T., HADNOTT, L. L., AND FISHMAN, A. (2003). The efficacy of silver-ion implanted catheters in reducing peritoneal dialysis-related infections. *Peritoneal Dialysis International*, **23** (4), 368–374.

CROWLEY, J. (1974). A note on some recent likelihoods leading to the log rank test. *Biometrika*, **61** (3), 533–538.

DAVIES, S., BRYAN, J., PHILLIPS, L., AND RUSSELL, G. (1996). Longitudinal changes in peritoneal kinetics: the effects of peritoneal dialysis and peritonitis. *Nephrology Dialysis Transplantation*, **11** (3), 498–506.

DOMBROS, N., DRATWA, M., GOKAL, R., HEIMBÜRGER, O., KREDIET, R., PLUM, J., RODRIGUES, A., SELGAS, R., STRUIJK, D., AND VERGER, C. (2005). European best practice guidelines for peritoneal dialysis. 3 peritoneal access. *Nephrology, dialysis, transplantation*, (20).

DOS SANTOS, R. P., KONKEWICZ, L. R., NAGEL, F. M., LISBOA, T., XAVIER, R. C., JACOBY, T., GASTAL, S. L., KUPLICH, N. M., PIRES, M. R., LOVATTO, C. G., ET AL. (2013). Changes in hand hygiene compliance after a multimodal intervention and seasonality variation. *American journal of infection control*, **41** (11), 1012–1016.

FABER, D. (2003). Fivefold reduction in peritonitis using a multifaceted continuous quality initiative program. *Advances in peritoneal dialysis*, **19**.

FAN, X., HUANG, R., WANG, J., YE, H., GUO, Q., YI, C., LIN, J., ZHOU, Q., SHAO, F., YU, X., ET AL. (2014). Risk factors for the first episode of peritonitis in southern chinese continuous ambulatory peritoneal dialysis patients. *PloS one*, **9** (9), e107485.

FARIAS, M. G., SOUCIE, J. M., MCCLELLAN, W., AND MITCH, W. E. (1994). Race and the risk of peritonitis: an analysis of factors associated with the initial episode. *Kidney international*, **46** (5), 1392–1396.

FENG, S., WANG, Y., QIU, B., WANG, Z., JIANG, L., ZHAN, Z., JIANG, S., AND SHEN, H. (2016). Impact of early-onset peritonitis on mortality and technique survival in peritoneal dialysis patients. *SpringerPlus*, **5** (1), 1676.

FIGUEIREDO, A., GOH, B.-L., JENKINS, S., JOHNSON, D. W., MACTIER, R., RAMALAKSHMI, S., SHRESTHA, B., STRUIJK, D., AND WILKIE, M. (2010). Clinical practice guidelines for peritoneal access. *Peritoneal Dialysis International*, **30** (4), 424–429.

FIGUEIREDO, A. E., DE MORAES, T. P., BERNARDINI, J., POLI-DE FIGUEIREDO, C. E., BARRETTI, P., OLANDOSKI, M., AND PECOITS-FILHO, R. (2014a). Impact of patient training patterns on peritonitis rates in a large national cohort study. *Nephrology Dialysis Transplantation*, **30** (1), 137–142.

FIGUEIREDO, A. E., KEHL, A. C. G., BOTTIN, S. T., AND DAMINELLI, W. G. (2014b). Is the distribution of microorganisms and peritonitis affected by seasonality in peritoneal dialysis? *Open Journal of Nephrology*, **4** (04), 146.

FIGUEIREDO, A. E., POLI-DE FIGUEIREDO, C. E., MENEGHETTI, F., LISE, G. A. P., DETOFOLI, C. C., AND SILVA, L. B. D. (2013). Peritonitis in patients on peritoneal dialysis: analysis of a single brazilian center based on the international society for peritoneal dialysis. *Jornal Brasileiro de Nefrologia*, **35** (3), 214–219.

FITRIANTO, A. AND JIIN, R. L. T. (2013). Several types of residuals in cox regression model: An emprical study. *International Journal of Math anal*, **7**, 45–54.

FONTAN, M. P., RODRÍGUEZ-CARMONA, A., GARCÍA-NAVEIRO, R., ROSALES, M., VILLAVERDE, P., AND VALDÉS, F. (2005). Peritonitis-related mortality in patients undergoing chronic peritoneal dialysis. *Peritoneal Dialysis International*, **25** (3), 274–284.

FOURTOUNAS, C. (2011). The present and the future of peritoneal dialysis. *Hippokratia*, **15** (Suppl 2), 15–20.

FOX, J. (2002). Cox proportional-hazards regression for survival data. *An R and S-PLUS companion to applied regression*, **2002**, 1–18.

GAFAR, M. M. (2013). *Factors affecting treatment outcomes in tuberculosis (TB) patients in the Limpopo Province, South Africa*. Ph.D. thesis, University of Limpopo (Turfloop Campus).

GAIL, M., KRICKEBERG, K., SAMET, J., TSIATIS, A., AND WONG, W. (2007). Statistics for biology and health.

GEORGE, B., SEALS, S., AND ABAN, I. (2014). Survival analysis and regression models. *Journal of Nuclear Cardiology*, **21** (4), 686–694.

GHARIBVAND, L. AND FERNANDEZ, G. (2008). Advanced statistical and graphical features of sas® phreg. *In SAS Global Forum*. pp. 375–2008.

GHARIBVAND, L., JESKE, D. R., AND LIAO, S. (2008). Evaluation of a hospice care referral program using cox proportional hazards model. *In Western Users of SAS Software Conference, Universal City, CA*.

GOEL, M., KHANNA, P., AND KISHORE, J. (2010). Understanding survival analysis: Kaplan-meier estimate. *International journal of Ayurveda research*, **1** (4), 274.

GOKAL, R. AND MALLICK, N. (1999). Peritoneal dialysis. *The Lancet*, **353** (9155), 823–828.

GRAY, N. A., GRACE, B. S., AND MCDONALD, S. P. (2013). Peritoneal dialysis in rural Australia. *BMC nephrology*, **14** (1), 1–8.

HAN, S. H., LEE, S. C., AHN, S. V., LEE, J. E., KIM, D. K., LEE, T. H., MOON, S. J., KIM, B. S., KANG, S.-W., CHOI, K. H., ET AL. (2007). Reduced residual renal function is a risk of peritonitis in continuous ambulatory peritoneal dialysis patients. *Nephrology Dialysis Transplantation*, **22** (9), 2653–2658.

IGUMBOR, E. U. AND LAUBSCHER, D. B. R. (2003). Mortality profile from registered deaths for Limpopo province, South Africa.

IKABU, A., ASSOUNGA, A., AND APALATA, T. (2016a). A study of peritonitis in capd patients at Inkosi Albert Luthuli Central Hospital, Durban, South Africa. *J. innov. res.. health sci. biotechnol*, **1** (3), 65–76.

IKABU, A. S., ASSOUNGA, A. G. H., AND APALATA, T. (2016b). A study of peritonitis in continuous ambulatory peritoneal dialysis patients at Inkosi Albert Luthuli Central Hospital, Durban, South Africa.

ISLA, R. A. T., AMEH, O. I., MAPIYE, D., SWANEPOEL, C. R., BELLO, A. K., RATSELA, A. R., AND OKPECHI, I. G. (2016). Baseline predictors of mortality among predominantly rural-dwelling end-stage renal disease patients on chronic dialysis therapies in Limpopo, South Africa. *PloS one*, **11** (6), e0156642.

ISLA, R. A. T., MAPIYE, D., SWANEPOEL, C. R., ROZUMYK, N., HUBAHIB, J. E., AND OKPECHI, I. G. (2014). Continuous ambulatory peritoneal dialysis in Limpopo province, South Africa: predictors of patient and technique survival. *Peritoneal Dialysis International*, **34** (5), 518–525.

JAIN, D. C. AND VILCASSIM, N. J. (1991). Investigating household purchase timing decisions: A conditional hazard function approach. *Marketing Science*, **10** (1), 1–23.

JEDA CHINCHILLA, K. S., MELÉNDEZ, R., RAMAY, B., AND LOU-MEDA, R. (2016). Risk factors for peritonitis in a pediatric dialysis program in guatemala, 2011-2014.

JIEZHI, Q. (2009). *Comparison of Proportional Hazards and Accelerated Failure Time Models*. Ph.D. thesis, M. Sc. Thesis, University of Saskatchewan.

JOHNSON, D. W., HERZIG, K. A., PURDIE, D. M., CHANG, W., BROWN, A. M., RIGBY, R. J., CAMPBELL, S. B., NICOL, D. L., AND HAWLEY, C. M. (2000). Is

obesity a favorable prognostic factor in peritoneal dialysis patients? *Peritoneal dialysis international*, **20** (6), 715–721.

KARTAL, O. O. (2015). *Using survival analysis to investigate the persistence of students in an introductory information technology course at METU*. Ph.D. thesis, Middle East Technical University.

KATZ, I. J., SOFIANOU, L., AND HOPLEY, M. (2001). An african community-based chronic ambulatory peritoneal dialysis programme. *Nephrology Dialysis Transplantation*, **16** (12), 2395–2400.

KELECŞ, M., CETINKAYA, R., UYANIK, A., ACEMOUGLU, H., SAATÇI, F., AND UYANIK, M. H. (2010). Peritoneal dialysis-related peritonitis: an analysis of risk factors in northeast anatolia. *Turkish Journal of Medical Sciences*, **40** (4), 643–650.

KERSCHBAUM, J., KÖNIG, P., AND RUDNICKI, M. (2012). Risk factors associated with peritoneal-dialysis-related peritonitis. *International journal of nephrology*, **2012**.

KHANAL, S. P., SREENIVAS, V., AND ACHARYA, S. K. (2014). Accelerated failure time models: an application in the survival of acute liver failure patients in india. *Int. J. Sci. Res*, **3** (6), 161–166.

KIM, H. J., LEE, J., PARK, M., KIM, Y., LEE, H., KIM, D. K., JOO, K. W., KIM, Y. S., CHO, E. J., AHN, C., ET AL. (2017). Lower education level is a risk factor for peritonitis and technique failure but not a risk for overall mortality in peritoneal dialysis under comprehensive training system. *PloS one*, **12** (1), e0169063.

KLEIN, J. P., VAN HOUWELINGEN, H. C., IBRAHIM, J. G., AND SCHEIKE, T. H. (2013). *Handbook of survival analysis*. Chapman and Hall/CRC.

KORBET, S. M., VONESH, E. F., AND FIRANEK, C. A. (1993). A retrospective

assessment of risk factors for peritonitis among an urban capd population. *Peritoneal dialysis international*, **13** (2), 126–131.

KOTSANAS, D., POLKINGHORNE, K. R., KORMAN, T. M., ATKINS, R. C., AND BROWN, F. (2007). Risk factors for peritoneal dialysis-related peritonitis: Can we reduce the incidence and improve patient selection? *Nephrology*, **12** (3), 239–245.

KRISHNAN, M., THODIS, E., IKONOMOPOULOS, D., VIDGEN, E., CHU, M., BARGMAN, J. M., VAS, S. I., AND OREOPOULOS, D. G. (2002). Predictors of outcome following bacterial peritonitis in peritoneal dialysis. *Peritoneal Dialysis International*, **22** (5), 573–581.

KUMAR, V. A., SIDELL, M. A., YANG, W.-T., AND JONES, J. P. (2014). Predictors of peritonitis, hospital days, and technique survival for peritoneal dialysis patients in a managed care setting. *Peritoneal Dialysis International*, **34** (2), 171–178.

LI, P. K.-T., LAW, M. C., CHOW, K. M., LEUNG, C.-B., KWAN, B. C.-H., CHUNG, K. Y., AND SZETO, C.-C. (2007). Good patient and technique survival in elderly patients on continuous ambulatory peritoneal dialysis. *Peritoneal Dialysis International*, **27** (Supplement 2), S196–S201.

LI, P. K.-T., SZETO, C. C., PIRAINO, B., BERNARDINI, J., FIGUEIREDO, A. E., GUPTA, A., JOHNSON, D. W., KUIJPER, E. J., LYE, W.-C., SALZER, W., ET AL. (2010). Peritoneal dialysis-related infections recommendations: 2010 update. *Peritoneal Dialysis International*, **30** (4), 393–423.

LIM, W. H., BOUDVILLE, N., MCDONALD, S. P., GORHAM, G., JOHNSON, D. W., AND JOSE, M. (2011). Remote indigenous peritoneal dialysis patients have higher risk of peritonitis, technique failure, all-cause and peritonitis-related mortality. *Nephrology Dialysis Transplantation*, **26** (10), 3366–3372.

LIN, D. Y., WEI, L.-J., AND YING, Z. (1993). Checking the cox model with cumulative sums of martingale-based residuals. *Biometrika*, **80** (3), 557–572.

LIU, X. (2012). *Survival analysis: models and applications*. John Wiley & Sons.

LOBO, J. V. D., VILLAR, K. R., ANDRADE JÚNIOR, M. P. D., AND BASTOS, K. D. A. (2010). Predictor factors of peritoneal dialysis-related peritonitis. *Jornal Brasileiro de Nefrologia*, **32** (2), 156–164.

LONERGAN, M. (2013). Treatment of peritoneal dialysis–associated fungal peritonitis. *BMC nephrology*, 1–23.

LOYNES, R. (1969). On cox and snell's general definition of residuals. *Journal of the Royal Statistical Society. Series B (Methodological)*, 103–106.

MARTÍN, E. V. AND FERNÁNDEZ, E. (2006). Patients with learning difficulties: outcome on peritoneal dialysis. *Advances in Peritoneal Dialysis*, **22**, 116–118.

MARTIN, L. C., CARAMORI, J. C., FERNANDES, N., DIVINO-FILHO, J. C., PECOITS-FILHO, R., AND BARRETTI, P. (2011). Geographic and educational factors and risk of the first peritonitis episode in brazilian peritoneal dialysis study (brazpd) patients. *Clinical Journal of the American Society of Nephrology*, **6** (8), 1944–1951.

MCDONALD, S. P., COLLINS, J. F., RUMPSFELD, M., AND JOHNSON, D. W. (2004). Obesity is a risk factor for peritonitis in the australian and new zealand peritoneal dialysis patient populations. *Peritoneal Dialysis International*, **24** (4), 340–346.

MEHROTRA, R. AND BOESCHOTEN, E. (2009). Current status of peritoneal dialysis. *In Nolph and Gokal's Textbook of Peritoneal Dialysis*. Springer, pp. 19–37.

MILLER JR, R. G. (1983). What price kaplan-meier? *Biometrics*, 1077–1081.

MILLER JR, R. G. (2011). *Survival analysis*, volume 66. John Wiley & Sons.

MONCRIEF, J. W. AND POPOVICH, R. P. (1979). Continuous ambulatory peritoneal dialysis1. *In Today's Art of Peritoneal Dialysis*, volume 17. Karger Publishers, pp. 139–145.

MUSTAFA, M., MENON, J., HAMED, S., ILLZAM, E., SHARIFA, A., AND SHM, A. (2015). Diagnosis and treatment of peritonitis in patients on continuous ambulatory peritoneal dialysis. *IOSR Journal of Dental and Medical Sciences (IOSR-JDMS)*, **1** (14), 63–67.

NAYAK, K., PRABHU, M., SINOJ, K., SUBHRAMANYAM, S., AND SRIDHAR, G. (2009). Peritoneal dialysis in developing countries. *In Peritoneal Dialysis-From Basic Concepts to Clinical Excellence*, volume 163. Karger Publishers, pp. 270–277.

NEGOI, D. AND NOLPH, K. (2009). History of peritoneal dialysis. *In Nolph and Gokal's Textbook of Peritoneal Dialysis*. Springer, pp. 1–18.

NESSIM, S. J., BARGMAN, J. M., AUSTIN, P. C., STORY, K., AND JASSAL, S. V. (2009). Impact of age on peritonitis risk in peritoneal dialysis patients: an era effect. *Clinical Journal of the American Society of Nephrology*, **4** (1), 135–141.

NIETO-RÍOS, J. F., DÍAZ-BETANCUR, J. S., ARBELÁEZ-GÓMEZ, M., GARCÍA-GARCÍA, Á., RODELO-CEBALLOS, J., REINO-BUELVAS, A., SERNA-HIGUITA, L. M., AND HENAO-SIERRA, J. E. (2014). Peritoneal dialysis-related peritonitis: twenty-seven years of experience in a colombian medical center. *Nefrologia*, **34** (1), 88–95.

NISHINA, M., YANAGI, H., KAKUTA, T., ENDOH, M., FUKAGAWA, M., AND TAKAGI, A. (2014). A 10-year retrospective cohort study on the risk factors for peritoneal dialysis-related peritonitis: a single-center study at tokai university hospital. *Clinical and experimental nephrology*, **18** (4), 649.

OHNO-MACHADO, L. (2001). Modeling medical prognosis: survival analysis techniques. *Journal of biomedical informatics*, **34** (6), 428–439.

OKAYAMA, M., INOUE, T., NODAIRA, Y., KIMURA, Y., NOBE, K., SETO, T., SUEYOSHI, K., TAKANE, H., TAKENAKA, T., OKADA, H., ET AL. (2011). Aging is an important risk factor for peritoneal dialysis-associated peritonitis. *In*

Advances in peritoneal dialysis. Conference on Peritoneal Dialysis, volume 28. pp. 50–54.

OREOPOULOS, D. G. AND THODIS, E. (2010). The history of peritoneal dialysis: early years at toronto western hospital. *Dialysis & Transplantation*, **39** (8), 338–343.

PAUW, K. ET AL. (2005). A profile of the limpopo province: demographics, poverty, inequality, and unemployment. Technical report, Provide Project.

PERENCEVICH, E. N., MCGREGOR, J. C., SHARDELL, M., FURUNO, J. P., HARRIS, A. D., MORRIS, J. G., FISMAN, D. N., AND JOHNSON, J. A. (2008). Summer peaks in the incidences of gram-negative bacterial infection among hospitalized patients. *Infection Control & Hospital Epidemiology*, **29** (12), 1124–1131.

PETERS, A. (2014). Safety issues in home dialysis. *Nephrology Nursing Journal*, **41** (1), 89.

PETO, P. AND PIKE, M. (1977). Conservation of the approximation 2/e in logrank-test survival data on tumour incidence data. *Biometrika*, **29**, 579.

PIRAINO, B. (1998). Diseaseof the month. *J Am Soc Nephrol*, **9**, 1956–1964.

PIRAINO, B., BERNARDINI, J., BROWN, E., FIGUEIREDO, A., JOHNSON, D. W., LYE, W.-C., PRICE, V., RAMALAKSHMI, S., AND SZETO, C.-C. (2011). Ispd position statement on reducing the risks of peritoneal dialysis–related infections. *Peritoneal Dialysis International*, **31** (6), 614–630.

PIRAINO, B., BERNARDINI, J., CENTA, P. K., SORKIN, M., ET AL. (1991). The effect of body weight on capd related infections and catheter loss. *Peritoneal Dialysis International*, **11** (1), 64–68.

POLLOCK, K. H., WINTERSTEIN, S. R., BUNCK, C. M., AND CURTIS, P. D. (1989). Survival analysis in telemetry studies: the staggered entry design. *The Journal of Wildlife Management*, 7–15.

PRASAD, N., GUPTA, A., SHARMA, R. K., PRASAD, K. N., GULATI, S., AND SHARMA, A. P. (2003). Outcome of gram-positive and gram-negative peritonitis in patients on continuous ambulatory peritoneal dialysis: a single-center experience. *Peritoneal dialysis international*, **23** (Supplement 2), S144–S147.

PRASAD, N., SINHA, A., GUPTA, A., SHARMA, R. K., BHADAURIA, D., CHANDRA, A., PRASAD, K. N., AND KAUL, A. (2014). Effect of body mass index on outcomes of peritoneal dialysis patients in india. *Peritoneal Dialysis International*, **34** (4), 399–408.

RAAIJMAKERS, R., GAJJAR, P., SCHRÖDER, C., AND NOURSE, P. (2010). Peritonitis in children on peritoneal dialysis in cape town, south africa: epidemiology and risks. *Pediatric Nephrology*, **25** (10), 2149–2157.

RODRIGUEZ, G. (2005). Non-parametric estimation in survival models.

ROONEY, C. (2015). *Using survival analysis to identify the determinants of academic exclusion and graduation in three faculties at UCT*.

ROYSTON, P. AND PARMAR, M. K. (2013). Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. *BMC medical research methodology*, **13** (1), 1–15.

RUDNICKI, M., KERSCHBAUM, J., HAUSDORFER, J., MAYER, G., AND KÖNIG, P. (2010). Risk factors for peritoneal dialysis–associated peritonitis: the role of oral active vitamin d. *Peritoneal Dialysis International*, **30** (5), 541–548.

SANABRIA, M., DEVIA, M., HERNÁNDEZ, G., ASTUDILLO, K., TRILLOS, C., URIBE, M., LATORRE, C., BERNAL, A., RIVERA, A., ET AL. (2015). Outcomes of a peritoneal dialysis program in remote communities within colombia. *Peritoneal Dialysis International*, **35** (1), 52–61.

SANTHAKUMARAN, T., SAMAD, N., AND FAN, S. L. (2016). Hydration status measured by bcm: A potential modifiable risk factor for peritonitis in patients on peritoneal dialysis. *Nephrology*, **21** (5), 404–409.

SAYED, S. A., ABU-AISHA, H., AHMED, M. E., AND ELAMIN, S. (2013). Effect of the patient's knowledge on peritonitis rates in peritoneal dialysis. *Peritoneal Dialysis International*, **33** (4), 362–366.

STANLEY, C., MOLYNEUX, E., AND MUKAKA, M. (2016). Comparison of performance of exponential, cox proportional hazards, weibull and frailty survival models for analysis of small sample size data. *Journal of Medical Statistics and Informatics*, **4** (1), 2.

STASON, W. B., BARNES, B. A., AND DIRECTOR, N. E. O. B. (1985). The effectiveness and costs of continuous ambulatory peritoneal dialysis (capd). Congress of the US, Office of Technology Assessment.

SWAIN, P. K. AND GROVER, G. (2016). Determination of predictors associated with hiv/aids patients on art using accelerated failure time model for interval censored survival data. *American Journal of Biostatistics*.

TAHA, E. T. A. ET AL. (2017). *Dialysis Solution Analysis for Peritonitis Infection Prevention by using Optical System Prototype*. Ph.D. thesis, Sudan University of science and technology.

TRAVAR, M., VLATKOVIC, V., AND VOJVODIC, D. (2015). Microbiological aspects of peritonitis in patients on continuous ambulatory peritoneal dialysis: A monocentric five years follow up study. *Journal of Infectious Diseases & Therapy*.

TSUBAKIHARA, Y., NISHI, S., AKIBA, T., HIRAKATA, H., ISEKI, K., KUBOTA, M., KURIYAMA, S., KOMATSU, Y., SUZUKI, M., NAKAI, S., ET AL. (2010). 2008 japanese society for dialysis therapy: guidelines for renal anemia in chronic kidney disease. *Therapeutic Apheresis and Dialysis*, **14** (3), 240–275.

TWARDOWSKI, Z. J. AND PROWANT, B. F. (1996). Exit-site healing post catheter implantation. *Peritoneal dialysis international*, **16** (Suppl 3), S51–S70.

VAN, S. ET AL. (2014). *Effects of inflammation and infection on peritoneal transport*. Boxpress.

VAN ESCH, S., STRUIJK, D. G., AND KREDIET, R. T. (2016). The natural time course of membrane alterations during peritoneal dialysis is partly altered by peritonitis. *Peritoneal Dialysis International*, **36** (4), 448–456.

WANG, Z., ZHANG, Y., XIONG, F., LI, H., DING, Y., GAO, Y., ZHAO, L., AND WAN, S. (2015). Association between medical insurance type and survival in patients undergoing peritoneal dialysis. *BMC nephrology*, **16** (1), 33.

WINTERBOTTOM, J. (2015). *A cross-sectional, correlational survey to explore the relationship between Renal Association biochemical and haematological markers and health-related quality of life in patients receiving haemodialysis in the North West of England*. Ph.D. thesis, University of Manchester.

WONG, J., TALJAARD, M., FORSTER, A. J., ESCOBAR, G. J., AND VAN WALRAVEN, C. (2011). Derivation and validation of a model to predict daily risk of death in hospital. *Medical care*, **49** (8), 734–743.

XIN, X. (2011). *A study of ties and time-varying covariates in cox proportional hazards model*. Ph.D. thesis, University of Guelph.

XU, R., HAN, Q.-F., ZHU, T.-Y., REN, Y.-P., CHEN, J.-H., ZHAO, H.-P., CHEN, M.-H., DONG, J., WANG, Y., HAO, C.-M., ET AL. (2012). Impact of individual and environmental socioeconomic status on peritoneal dialysis outcomes: a retrospective multicenter cohort study. *PloS one*, **7** (11), e50766.

ZHAOJIN, C. (2012). *Handling of Tied Failures in Competing Risks Analysis*. Ph.D. thesis, National University of Singapore.