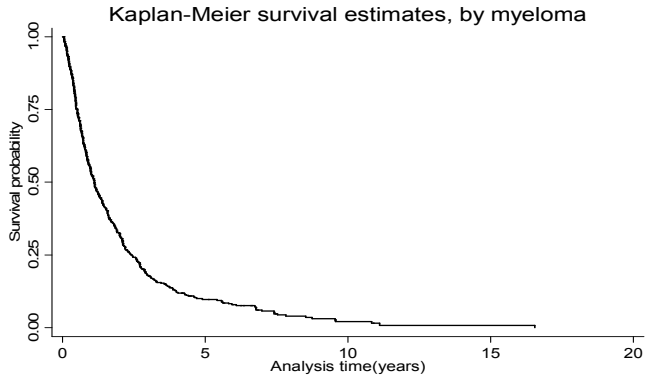
Biostatistics Collaboration of Australia

Work Placement Project

Survival of Myeloma Patients on Dialysis



Fakhrul Islam
Supervisor: Dr Patrick Kelly
Co-Supervisor: Dr Angela Webster
School of Public Health

University of Sydney

Table of Contents

Preface

Introduction

The project titled "Survival of Myeloma Patients on dialysis" was carried out under the supervision of Dr Patrick Kelly, a senior Biostatistician of the School of Public Health and with the co-supervision of Dr Angela Webster, a Nephrologist and Epidemiologist of the same school. The aim of the project was to describe the survival of Myeloma; the second most commonly diagnosed blood cancer in Australia.

Student's role

My role was to prepare the data for analysis, advising the appropriate analysis, implement the agreed analysis plan and interpret the final results.

The project involved much data manipulation and the majority of my time was spent on this. Although I was primarily responsible for analysing the data, I could consult with my supervisor and co-supervisor to clarify any statistical issues.

Reflection on learning

Communication skills:

I met with my supervisor weekly. During those meetings I explained the work I had conducted and presented my results. We also communicated via email – sending in advance my results and summarising our meetings. I also learnt how to write a report after analysing the data. Thus, the project enhanced my communication skills both verbally and written, while I was presenting my reports and discussed the results on a regular basis with my supervisors.

Work patterns/planning:

The project developed my planning and organisation skills as well. To implement the project, I met with my supervisor in regular meetings, contacted through email, sent summaries of analysis via email and so on. Also, I maintained a project diary that included the meeting agenda and listing of post meeting outcomes. Thus, I was able to integrate the results of the whole project. As a result, I learnt how a project is planned and managed within a time frame. I was always very prompt in replying to my supervisor's emails and always attended the meetings on time and well prepared.

Statistical principles:

The project also increased my applied statistical knowledge and skills through hypothesis testing and making inference on the test result. More specifically, I had to develop a number of models identifying the significant covariates associated with the risk factors of the survival of myeloma patients and estimate the risk of death for myeloma patients compared to non-myeloma dialysis patients who reached the End Stage Kidney Disease (ESKD) stage.

Statistical issues

The statistical issues involved the appropriate selection of methods. The proposed methods were the survival analysis of the censored data and choosing the correct model including selection of appropriate covariates for the model.

Statistical methods:

The statistical analysis involved methods of applying Kaplan-Meier curves, Cox proportional hazard models and Stratified Cox models, which I learnt specifically from 'Survival Analysis'. Also, I applied my learnt knowledge of 'Categorical Data Analysis' in order to convert continuous variables to categorical variables. I also referred to various statistics books, such as Applied Survival Analysis: Regression modelling of time to event data by Hosmer DW. &

Lemeshow S. (1999), Survival Analysis: A self-learning text by Kleinbaum DG. & Klein M. (1996) and so on.

Statistical computing:

The work placement project substantially increased my knowledge of statistical computing particularly in STATA and some extent SAS, as all analyses and data management were conducted in these packages. I also learnt a lot about manipulation and data cleaning. For example, each patient had several rows of data and had to be compressed into a single row per patient data without losing relevant information for the analyses. I also generated a number of new variables, performed cross-tabulations, created various types of graphs and so on. The details of these data manipulation process have been explained more elaborately in the 'data manipulation' area.

Teamwork

Communication with other team members:

I had to communicate with other two team members, my supervisor, Dr Patrick Kelly, a senior Biostatistician of the School of Public Health and the co-supervisor, Dr Angela Webster, a Nephrologist and Epidemiologist of the same school. I met with them on weekly basis and sometimes through e-mail regarding progress of my work. As a result, I developed a sound team work skills.

Working within timelines:

I was able to present the results of the weekly proposed analysis and submitted the portfolio within a time frame as instructed by my supervisor. More specifically, I was given set tasks at each meeting such as data manipulation, analysis and achieved tasks set for the next meeting. I used to use a project diary, where I listed all those activities and maintain the diary accordingly. As a result, I have achieved how to work within a time frame.

Helping others to understand statistical issues-teaching:

The project team consists of three members, one of which was an Epidemiologist. Consequently, I had to explain the results to her and any statistical issues arose in the meeting particularly the models.

Ethical considerations

I signed the confidentiality agreement while I collected ANZDATA cancer registry data to protect privacy of the patients. The signed agreement form is attached in appendix.

Confidentiality issues

I was strictly undertaken to preserve the confidentiality of the data and not to disclose any information of the output. In addition, a copy of the project will be sent to ANZDATA when published.

Professional responsibility

The data were kept password protected to fulfil my undertaken as a professional statistician and the communications were limited to the project team members only.

Project

Project title

Survival of Myeloma patients on dialysis

Location and Dates

School of Public Health, University of Sydney

July 2008-January 2009

Context

Dr Angela Webster, the Project Manager and the Chair of the cancer working group of ANZDATA was responsible for advising the clinical investigation of the survey data, which are updated every after six months. As part of her work, I was given the responsibility of analysing the data to investigate the risk factors associated with survival for myeloma patients, who had reached end stage kidney disease.

Contribution of student

As part of my analysis process, I proposed Dr Judy Simpson regarding the details of my analysis plan with the consultation of Dr Patrick Kelly that such data can be analysed using Kaplan-Meier curves and Cox proportional Hazard models. As a result, I started to meet my supervisors on weekly basis to present the results of my progressive work. I carried out all the analysis with the consultation of Dr Patrick Kelly and Dr Angela Webster using STATA statistical software.

Acknowledgements

I am indebted to the staff of the ANZDATA registry and to the nephrologists and staff (past and present) of all Australian and New Zealand renal units for their efforts in regular data collection; the impressive resource that is the ANZDATA Registry exists because of their dedicated efforts

Student's declaration

I declare that this project is my own work, under the supervision of Dr Patrick Kelly and with the co-supervision of Dr Angela Webster, and it has not previously been submitted for academic credit.

_____

Md. Fakhrul Islam

Date: 18 March 2009

Supervisor's declaration

The work in this project has been Fakhrul's own work.

_____

Dr Patrick Kelly

Date: 18 March 2009

Analysis of End Stage Kidney Disease (ESKD) data using Kaplan-Meier survival curves and Cox Proportional Hazard models

1. Project description

End-stage kidney disease (ESKD) or End Stage Renal Disease (ESRD) is a complete or nearly complete failure of the kidneys to function to excrete wastes, concentrate urine, and regulate electrolytes. End-stage kidney disease occurs usually at the chronic stage. More specifically, ESKD occurs when kidney function reach at the point of lower than 10% of its normal function. The kidney function is so low at this stage is that a patient without a dialysis or transplantation will die from accumulation of fluids and waste products in the body (UMM, 2007).

ANZDATA is the Australian and New Zealand Dialysis and Transplant Registry that collects a wide range of statistics which relate to the outcomes of treatment of those with end stage renal failure. The ANZDATA is comprised of patients commencing from 1963, the year renal replacement therapy was first used in Australia to 2006, who have undergone dialysis and/or a kidney transplant for the treatment of ESKD. The survey data were designed through circulation of printed survey forms for each patient at six-month intervals to all dialysis and transplant units in Australia and New Zealand. As part of the routine medical care, the kidney specialists collect certain information from the people receiving treatment with dialysis or kidney transplantation.

1.1 Background, rationale for project

Myeloma, also known as Multiple Myeloma or plasma cell Myeloma, is a cancer of the blood in which malignant plasma cells are overproduced in the bone marrow. Myeloma cells typically produce excessive amounts of paraprotein or M protein and can cause kidney damage that may progress to End Stage Kidney Disease (ESKD). Multiple Myeloma is the second most commonly diagnosed blood cancer and around 1200 people are newly diagnosed each year in Australia and it will increase as the population ages. Almost 80% of People diagnosed with Myeloma are over 60 years old and is uncommon in people under 40 years and it also occurs more frequently in men than in women.

According to the Australian Institute of Health and Welfare (AIHW), more Australians than ever are developing Myeloma and Non-Hodgkin's lymphoma (NHL). In recently released statistics, the AIHW listed the two cancers among the top five cancers. The incidence of Myeloma had increased by 44% in the 10 years from 1993 to 2003. A total of 2378 new patients were diagnosed for end-stage renal failure in Australia in 2006 (ANZDATA), a rate of 115 per million population per year. People who have reached ESKD will require permanent renal replacement therapy either by means of dialysis or kidney transplantation to improve their quality of life. The standardised incidence ratio of Myeloma cancers (ICD: C90) of Australian ESKD patients on dialysis is 9.58(95%CI, 7.64-11.86) (JAMA, 2006). According to the British Medical Journal, dialysis is recommended for the treatment of myeloma, as adequate dialysis may improve patient tolerance to chemotherapy. It is also reported that the disease may be considered as incurable and thus fatal. As a result, renal transplantation for the myeloma patients is discouraged. According to the journal named Nephrology Dialysis Transplantation "fear of acceleration of myeloma following immunosuppressive therapy directed primarily at T-cells" (NDT, 1996).

1.2 Aim

The project has mainly focussed on the myeloma cancer patients who were treated with dialysis, to identify their survival characteristics. The exploratory data analysis has been conducted to identify the focussed groups. Cox proportional hazard models were developed for myeloma and non-myeloma patients by means of statistical software STATA.

Using data supplied by the Australia and New Zealand Dialysis and Transplant Data Registry (ANZDATA), this project will describe the survival of Myeloma patients on dialysis and identify the characteristics of improved survival. To identify the variables associated with improved survival, the data will be analysed using Kaplan-Meier survival curves, log-rank tests and Cox proportional hazards models.

The two main aims of the project were:

(i) Estimate the risk of death of myeloma patients compared to non-myeloma dialysis patients, adjusting for confounding factors.

(ii) Describe the survival and risk factors of survival for myeloma patients.

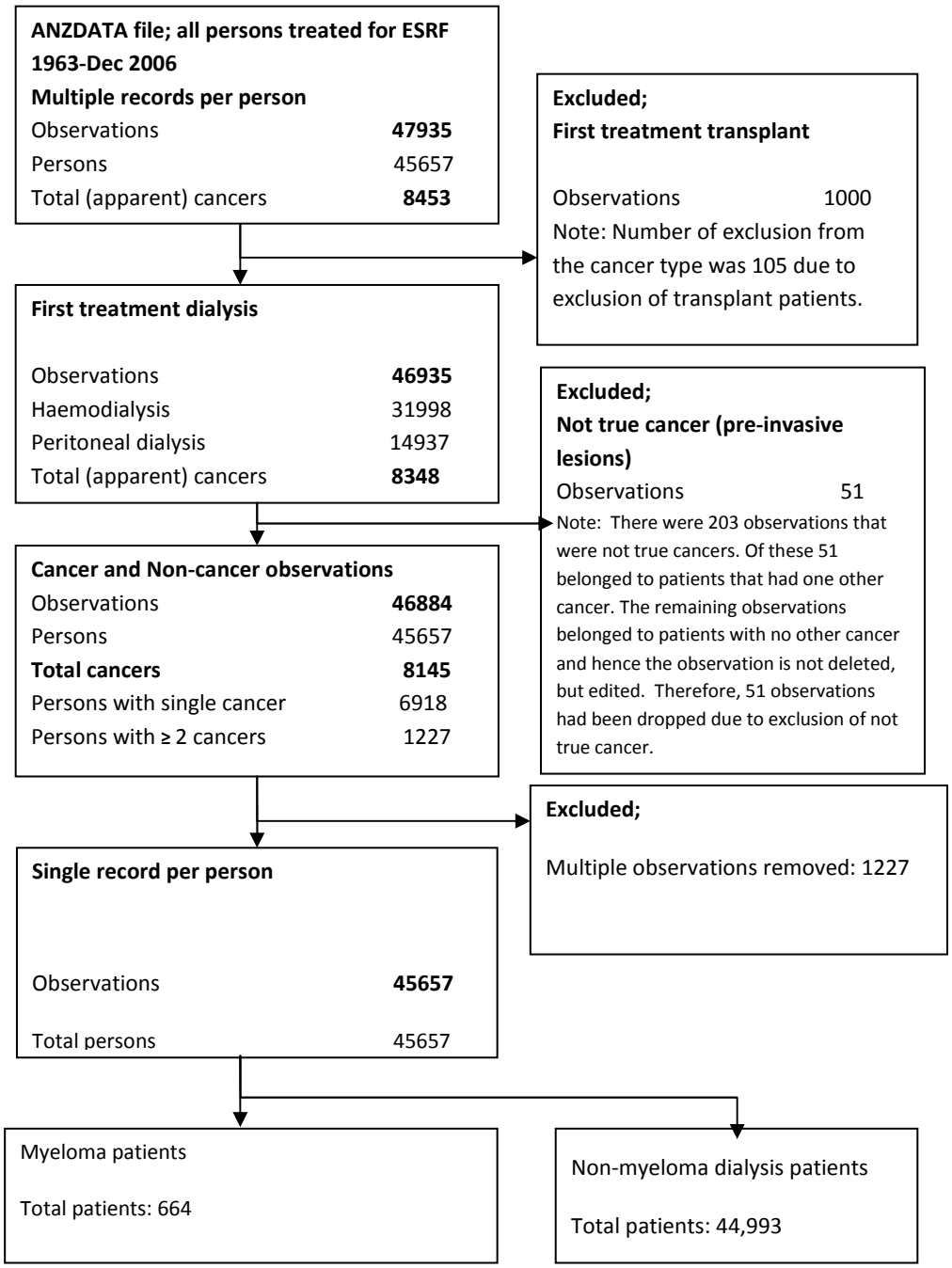
2. Data management

Data were supplied from Dr Angela Webster in Stata format, where each row corresponds to a cancer for an ESKD patient. If a patient had one row then they had either '0' or '1' cancer. If several rows correspond to a patient, then the number of rows is equal to the number of cancers.

The primary data set was comprised of 47935 multiple observations. I had to create a single observation file of 45657 patients from multiple observations that is one row per patient. First,

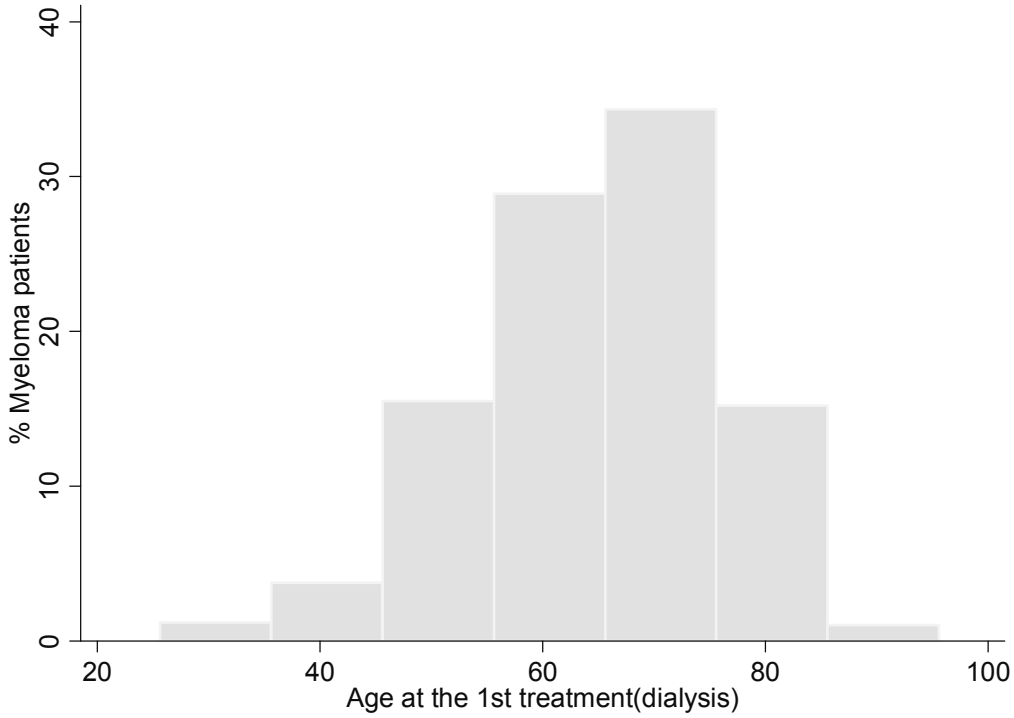
data were sorted out the cervix cancers from the Cervical Intraepithelial Neoplasia (CIN), a type of cervical pre cancers, which were coded as cancers. The total number of cancers found was 8453 out of which 212 were cervix cancers. Second, the dataset had two variables for identifying the type of cancer a patient had – "catype", which coded the different types of cancers a number and "typedecr" which was a was a string variable describing the type of cancer in words. I checked for any discrepancies between "catype" and "typedescr"- no error was found in the data set. Third, we were only interested in dialysis patients because myeloma patients are not eligible for a kidney transplant. Hence all transplant patients were removed – this only included one myeloma patient, who was diagnosed with myeloma many years after being diagnosed with EKSD. Fourth, I sorted the records for each patient- by cancer date by generating sequence variable. Fifth, I created new variables such as 'myeloma' as cancer type of 10. It was confirmed by counting and cross-tabulating myeloma and its sequences that myeloma could not be seen as the multiple occurrence of disease for the same patient. Another new variable 'canceroccureskd' was created from duration of myeloma and ESKD. The details of the new variables and categorisation of those variables can be seen in Table-1 in the statistical methods section. Finally, the records for each patient were compressed to a single row, but keeping such information such the number of total cancers. The details of the steps are summarised by means of the following flow chart (Figure 1). Out of the 45657 patients, 664 were patients who had myeloma.

Figure 1: Flow chart from multiple records per person (1 record per cancer) of all dialysis patients in ANZDATA, to single record per patient for those having had at least one myeloma where >1 cancer had occurred.



I constructed histograms and tables to check the variables such as age, gender, time between myeloma and ESKD and so on. See for example, Figure 2, which shows myeloma was very uncommon under the age of 40.

Figure 2: Histogram of age at the first treatment of dialysis for myeloma patients.



3. Statistical Methods

3.1 Exploratory Data Analysis

Table 1 summarises the characteristics of both myeloma and non-myeloma patients.

Table 1: Characteristics of dialysis population with and without myeloma, in Australia and New Zealand, 1963-2006

Characteristic	Myeloma patients		Other patients	
Total	**(n=664)**	**%**	**(n=44993)**	**%**
Age at ESKD*				
<60	210	31.6	26946	59.9
60-70	209	31.5	9898	22.0
>=70	245	36.9	8149	18.1
Gender				
Female	251	37.8	19186	42.6
Male	413	62.2	25807	57.4
Dialysis modality				
haemodialysis	533	80.3	30477	67.7
Peritoneal dialysis	131	19.7	14516	32.3
Diagnosis of ESKD				
before 1996	208	31.3	22717	50.5
1996-2002	214	32.2	12345	27.4
2002-present	242	36.5	9931	22.1
Primary renal disease				
Myeloma	488	73.5	4**	0.0
Possibly myeloma related	54	8.1	928	2.1
Glomerulonephritis/ IgA nephropathy	44	6.6	14073	31.3
Other causes	78	11.8	29,988	66.6
Racial background				
Non-white	40	6.0	9348	20.9
White	624	94.0	35609	79.1

Other malignancy prior to ESKD				
None	298	44.9	42244	93.9
Pre-dialysis malignancy	366	55.1	2749	6.1
Other malignancy subsequent to ESKD				
None	563	84.8	2964	6.6
Post-dialysis malignancy	101	15.2	42029	93.4
Other malignancy at same time of ESKD				
None	660	99.4	44984	99.98
Other malignancy at dialysis	4	0.6	9	0.02
Status during follow-up				
Alive	126	19.0	18736	41.6
Died	538	81.0	26257	58.4
Smoking history at ESKD				
Yes	267	42.2	17133	38.1
No	397	59.8	27860	61.9
Diabetes Mellitus				
Yes	194	29.2	21860	48.6
No	470	70.8	23133	51.4

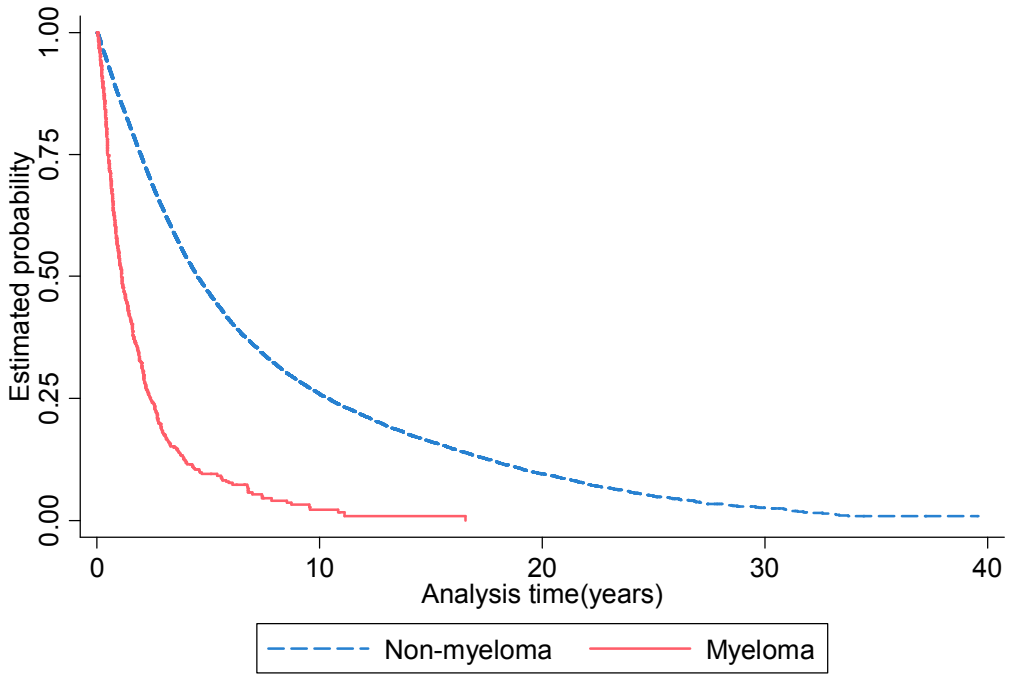
*End Stage Kidney Disease; ** Four non-myeloma patients suffered from Paraproteinaemia those included in Multiple myeloma group (Primary renal disease code 17)

We can see from the above table that myeloma is more common in men (62.2%) than in women. According to leukaemia foundation, it is found that myeloma is more common in men than in women (Leukaemia Foundation, 2006). Again, the result shows that the majority of people diagnosed with myeloma are over 60 years old (68.4%). The majority of myeloma patients are being treated by haemodialysis (80.3%), the remainder being treated by peritoneal dialysis. The report also revealed that most of primary renal diseases, within patients with myeloma, are

caused by myeloma (73.5%) disease where 79.4% patients are dying according to censoring status.

Figure 3 shows the survival of myeloma and non-myeloma patients. Time is measured from time of starting dialysis to death. Censoring occurs if the patient is either alive at last known date or else lost to follow-up. The median survival times of myeloma and non-myeloma patients are 1.1 and 4.5 years respectively.

Figure 3: Survival curves of myeloma and non-myeloma patients



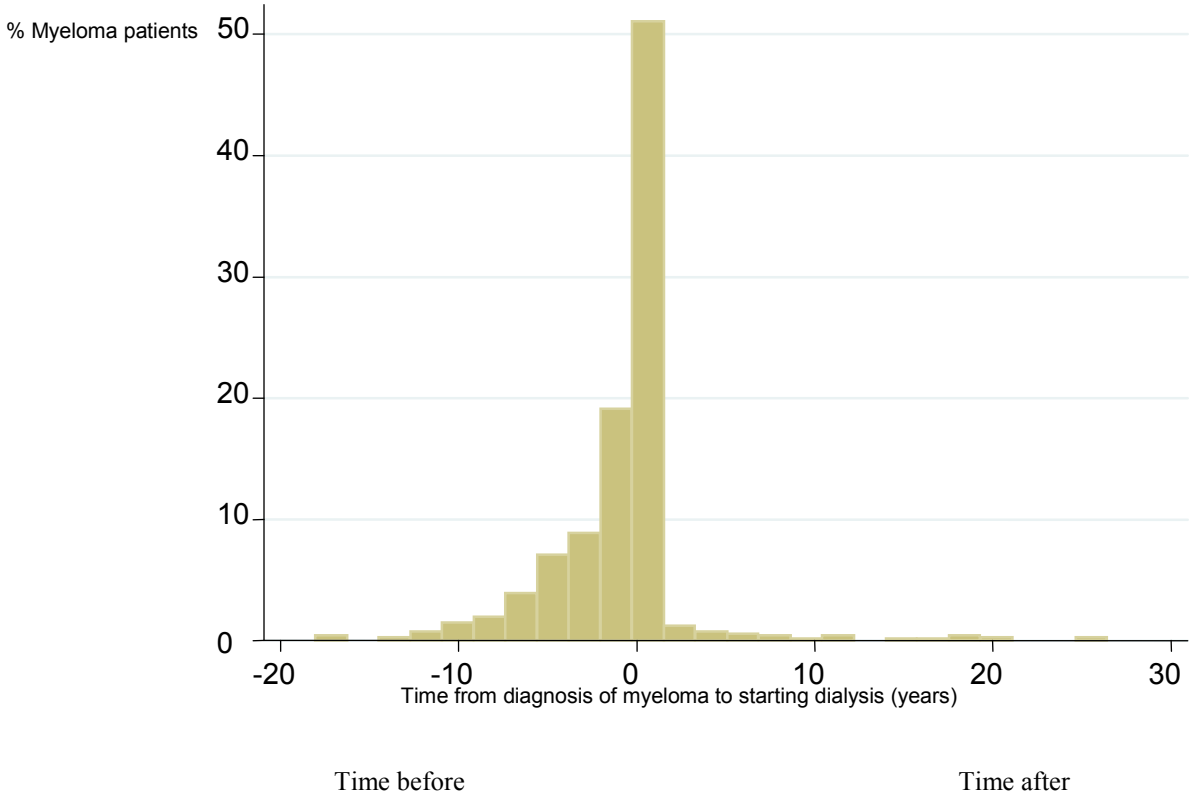
3.2 Analysing myeloma patients

The first aim of this project was to identify the risk factors that were associated with survival for myeloma patients. This section summarises the analysis that was conducted for this particular aim.

The total patients diagnosed with myeloma were 664. Figure 4 shows time from diagnosis of myeloma to starting diagnosis. A value of zero indicates those patients who were diagnosed with myeloma and starting dialysis at the same time (within 1 month). A value to the left of zero indicates those patients who were diagnosed with myeloma before ESKD and a value to the right indicates those patients who were diagnosed with myeloma after ESKD. Three hundred and sixty nine (369) patients (55.6%) were diagnosed with myeloma prior to starting dialysis, 193 (29.0%) of patients were diagnosed with myeloma within one month of starting dialysis and 102 (15.4%) of patients were diagnosed with myeloma after starting dialysis. Sixty one (61) patients (9.2%) were diagnosed with myeloma within one year of starting dialysis (ESKD) and 41 (6.2%) were diagnosed with myeloma more than 1 year after ESKD.

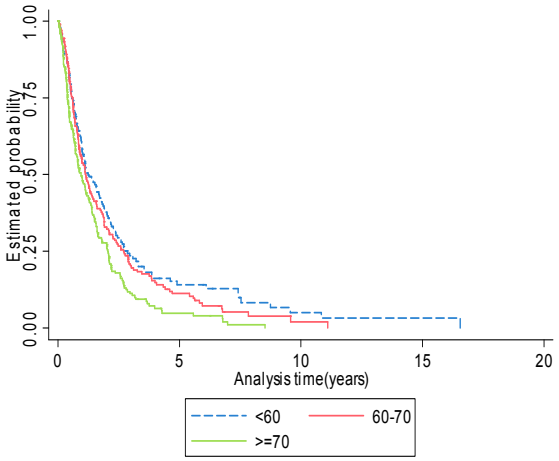
For this project, we wanted to determine the risk factors associated with survival for myeloma patients on dialysis. For this reason, it did not seem appropriate to include those patients that were on dialysis but did not have myeloma until a substantial time after being on dialysis. In consultation with Dr Webster, it was deemed plausible that myeloma may have been the cause of ESKD for patients who were diagnosed with myeloma for up to one year after beginning dialysis. Therefore, all patients who were diagnosed with myeloma more than one year after EKSD were excluded from the survival analysis- 41(6.2%) patients. Thus the results of the below analysis are based on 623 patients.

Figure 4: Histogram of time (years) between Myeloma and ESKD

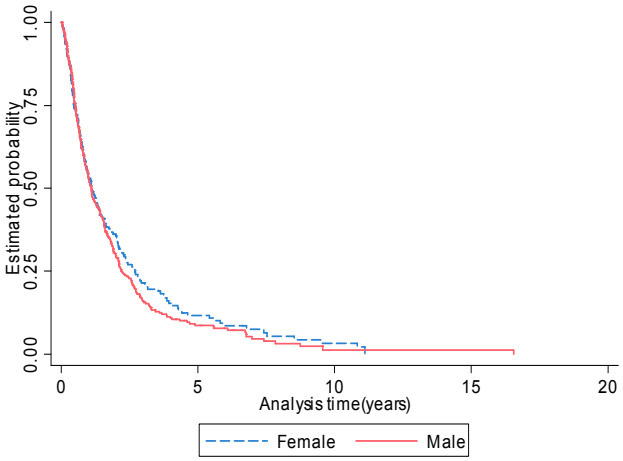


The survival analysis was with respect to time from ESKD to death. Of the 623, 117 (18.8%) were censored. Figure 5 shows the survival curves for potential risk factors of age, sex, treatment type, ethnicity, smoking status, diabetes mellitus, cause of primary kidney disease, year of diagnosis of ESKD, year of diagnosis of myeloma and duration between myeloma and ESKD. It is seen that there is no separation between the different levels for risk factors other than age, duration between myeloma and ESKD and primary cause of kidney disease.

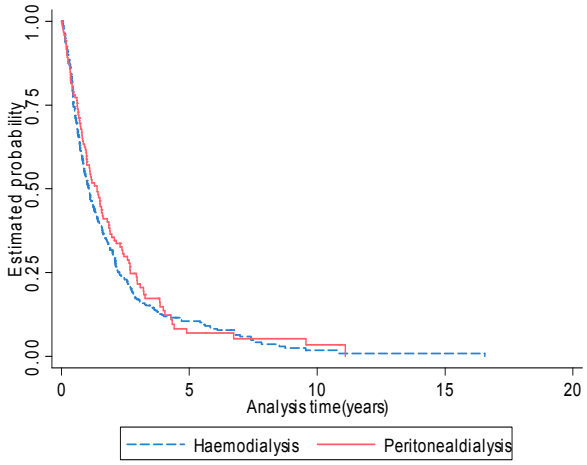
Figure 5: Kaplan-Meier survival curves of the covariates:



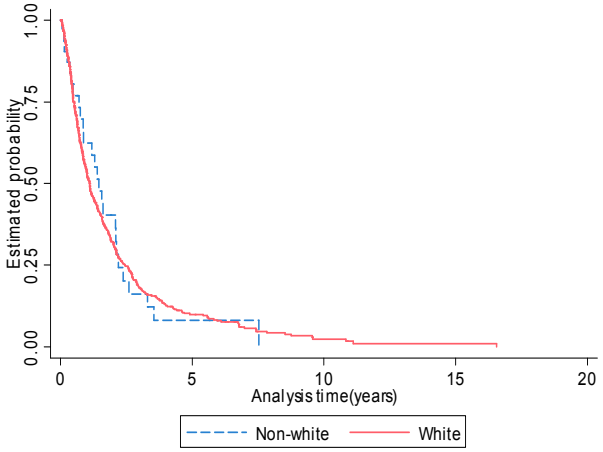
a) Probability of survival between different age groups



b) Probability of survival between male and female

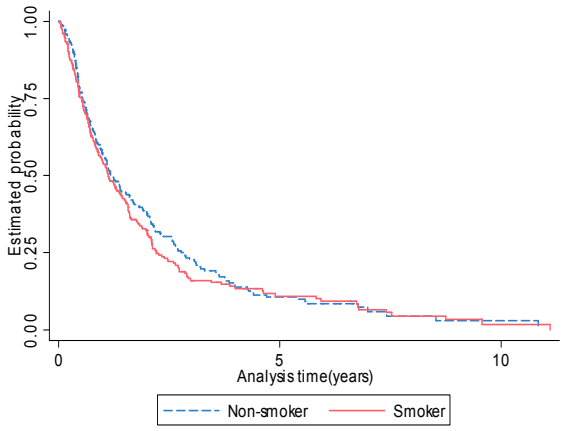


c) Probability of survival between haemodialysis and peritoneal dialysis

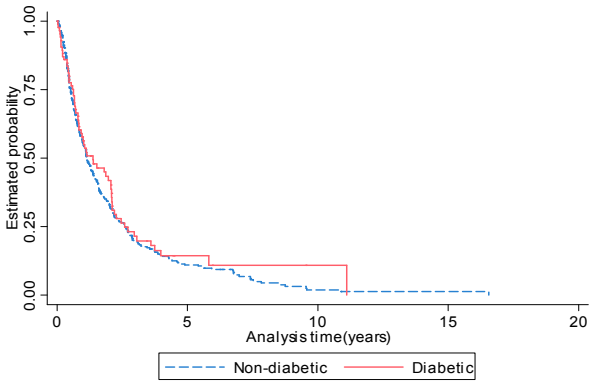


d) Probability of survival between ethnicity

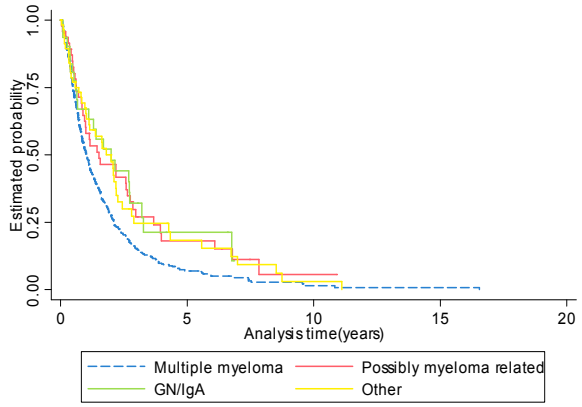
Figure 5 continued



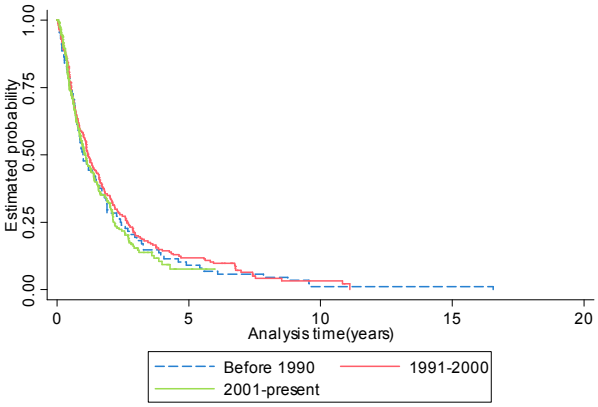
e) Probability of survival between smoker and non-smoker



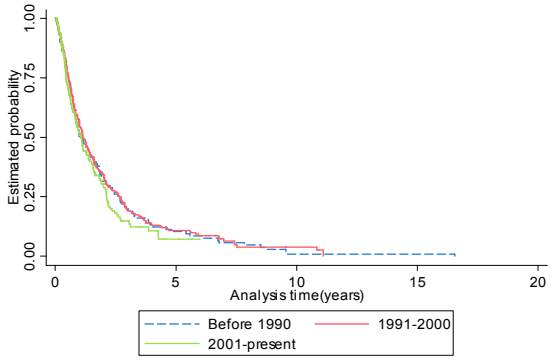
f) Probability of survival between diabetic and non-diabetic

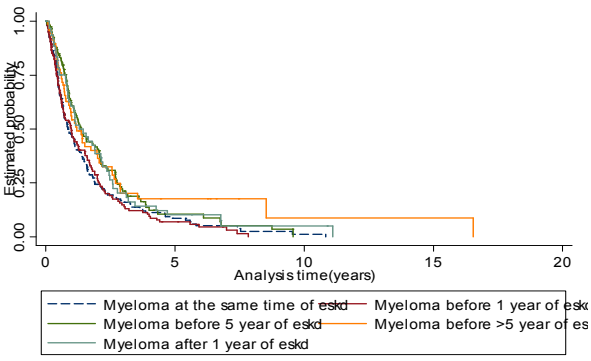


g) Probability of survival between different groups of primary kidney disease



h) Probability of survival between different group of year of diagnosis of ESKD





i) Probability of survival between different group of year of diagnosis of myeloma

j) Probability of survival between different group of Duration between myeloma and ESKD

Cox proportional hazard models were used for the analysis. The Cox model was first fitted singly to each variable. A multivariate model was then developed based on model building strategies proposed by Harrell (2006). Variables were entered into the model if the univariate $p< 0.25$. Variables were then dropped if $p >0.05$, with the exception of gender, which was included in the model regardless of statistical significance just to see the difference of hazard ratio between male and female. P-values for the Wald test are reported. The results of both the univariate and multivariate analysis are given in Table 3. Interaction terms were tested among age, myeloma and sex, but none was found to be statistically significant.

From the multivariate model, it can be seen that the only variables associated with survival were the age group. More specifically, the risk of dying of myeloma patients is 1.15 times more in patients aged 60 to 70 compared to those who are younger than 60 years. The risk of dying of myeloma patients is 1.69 times more in patients aged more than 70 compared to those who are younger than 60 years. Therefore, the risk of dying is increasing with increasing age.

The risk of dying is 33% lower in the possibly myeloma related disease compared to patients with multiple myeloma. The risk of dying is 40% lower in the GN/IgA group compared to patients with multiple myeloma. The risk of dying is 36% lower in other diseases group compared to patients with multiple myeloma. Therefore, if the patient's cause of ESKD was not diagnosed as myeloma, the risk of dying was lower.

The model also revealed that the risk of dying is 1.12 times more in male compared to female, which is not statistically significant ($p=0.23$).

The risk of dying of the patients who got myeloma after one year of ESKD is 0.85 times that of the patients who got myeloma at the same time of ESKD. The risk of dying of the patients who got myeloma before one year of ESKD is 1.02 times that of the patients who got myeloma at the same time of ESKD. The risk of dying of the patients who got myeloma before five year of ESKD is 0.74 times that of the patients who got myeloma at the same time of ESKD. The risk of

dying of the patients who got myeloma before greater than five years of ESKD is 0.73 times that of the patients who got myeloma at the same time of ESKD. Therefore, it was found that the risk of dying of the patients who got myeloma after ESKD is more than the patients who got myeloma before ESKD. Results seem to suggest that survival is slightly better amongst people who are diagnosed with myeloma more than one year before getting ESKD.

Table 3: Hazard ratios of death of myeloma patients on dialysis

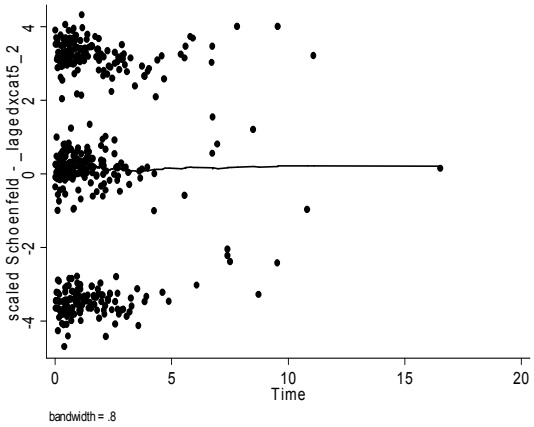
Covariates	**Hazard ratio**	**Univariate 95% CI**	**P-values**	**Hazard ratio**	**Multivariate 95% CI**	**P-values**
Age			0.001			<0.001
<60(referent)	-			-		
60-70	1.13	(0.91, 1.42)		1.15	(0.92, 1.44)	
>=70	1.51	(1.21, 1.88)		1.69	(1.35, 2.11)	
Sex			0.33			0.23
Female (referent)	-			-		
Male	1.09	(0.91, 1.31)		1.12	(0.93, 1.34)	
Primary kidney disease			0.002			0.003
Multiple myeloma(referent)	-			-		
Possibly myeloma related	0.65	(0.47, 0.91)		0.67	(0.47, 0.94)	
GN/IgA	0.62	(0.40, 0.95)		0.60	(0.39, 0.94)	
Other	0.68	(0.50, 0.93)		0.64	(0.46, 0.88)	
Duration between myeloma and ESKD			0.008			0.05
Myeloma before 1 year of ESKD	1.01	(0.80, 1.27)		1.02	(0.81, 1.29)	
Myeloma before 1 to 5 years of ESKD	0.72	(0.57, 0.92)		0.74	(0.58, 0.95)	
Myeloma before more than 5 years of ESKD	0.68	(0.50, 0.94)		0.73	(0.53, 1.01)	
Myeloma at the same time of ESKD (referent)						
Myeloma after 1 year of ESKD	0.74	(0.54, 1.02)		0.85	(0.61, 1.18)	

The proportional hazard (PH) assumption was checked for the final model. Stata tests for PH using global test. If the global test is not significant ($p>0.05$), it is assumed that the model satisfies the PH assumption. Table 4 shows that the global test was not significant ($p=0.52$). This indicated that the proportional hazard assumption was met. Also, it was found that each covariate met the PH assumption. We also looked at the graphical test of PH assumption behind this test (Figure 6) to see whether the smoothed lines are horizontal and centred at zero. The plots show that all the fitted lines are very flat indicating that the PH assumption was met.

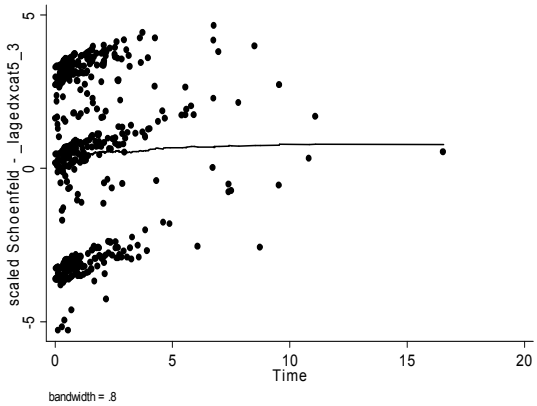
Table 4: Test of proportional hazards assumption

Variables	*Correlation coefficient (Rho)*	*Chi-squares*	*Degrees of freedom*	*P-values*
Age group				
<60 (referent)				
60-70	0.001	0.63	1	0.43
>70	-0.007	1.68	1	0.19
Sex				
Female (referent)				
Male	0.03	0.45	1	0.50
Primary kidney disease types				
Multiple Myeloma (referent)				
Possibly myeloma related	0.000	0.00	1	0.99
GN/IgA	-0.001	0.00	1	0.99
Other	0.03	0.50	1	0.48
Duration of Myeloma diagnosis				
Myeloma before 1 year of ESKD	0.01	0.07	1	0.80
Myeloma before 5 year of ESKD	0.03	0.57	1	0.45
Myeloma before >5 year of ESKD	-0.06	2.04	1	0.15
Myeloma at the same time of ESKD (referent)				
Myeloma after 1 year of ESKD	-0.003	0.00	1	0.95
Global Test		9.12	10	0.52

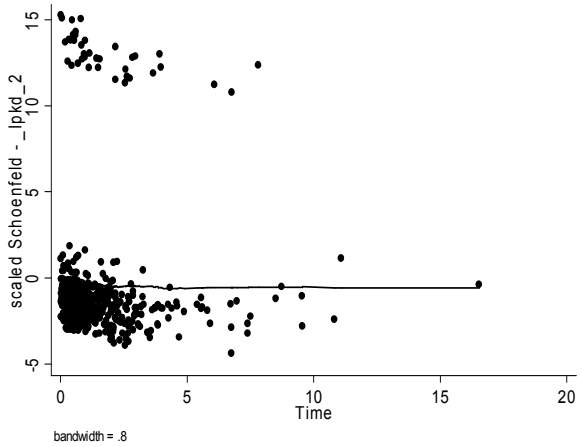
Figure 6: Graphical test of proportional hazard assumption



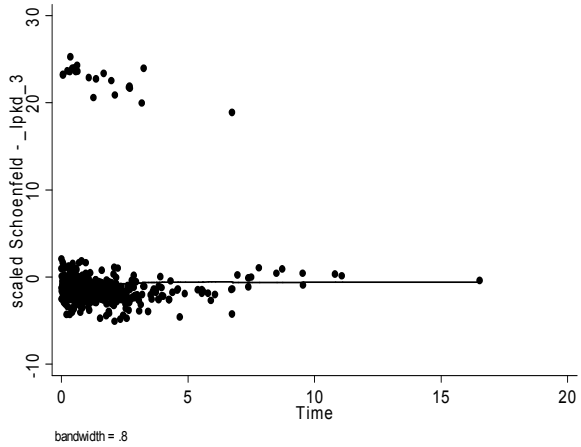
a) age group 60-70



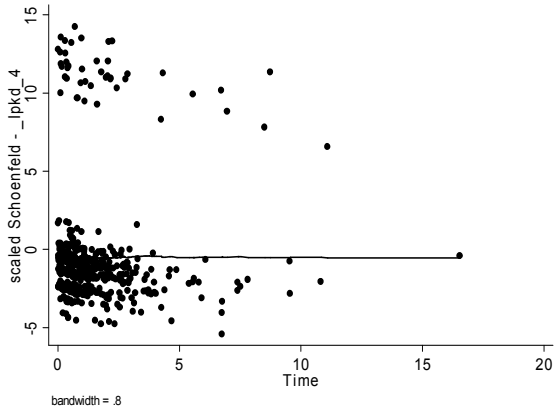
b) age group >70



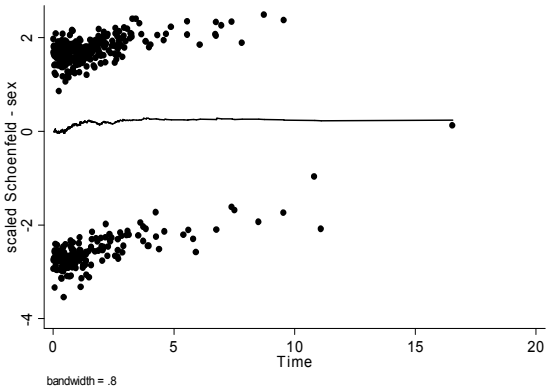
c) PKD (possibly myeloma related)



d) PKD (GN/IgA)



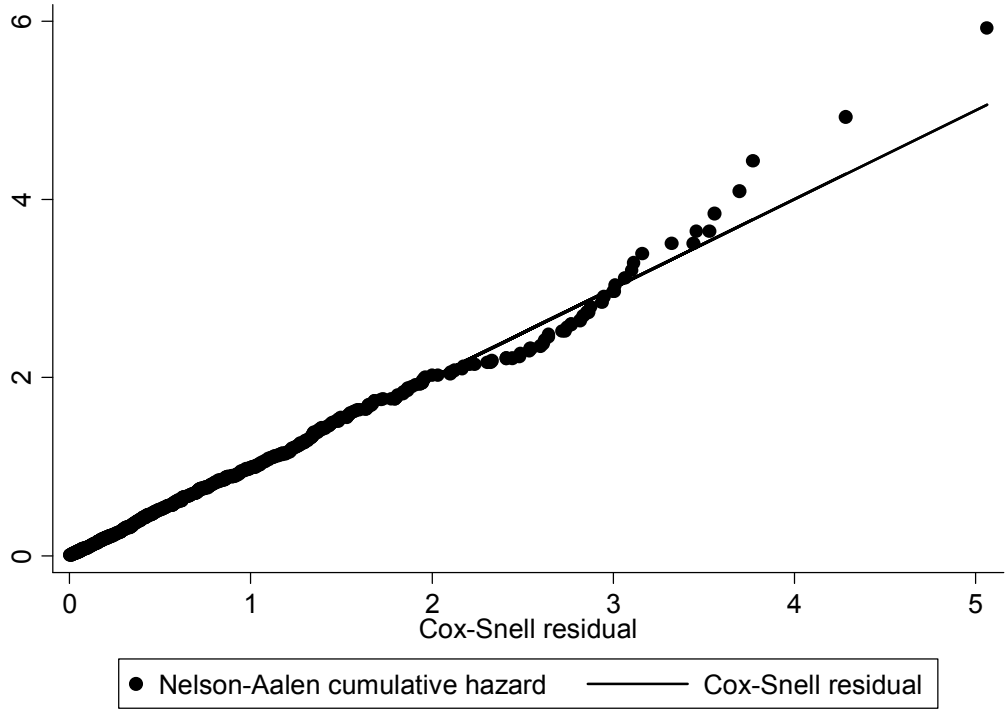
e) PKD (other)



f) sex (male)

The Cox-Snell residual goodness fit to test was used to determine whether the model fits the data well (p.225 Hosmer and Lemeshow, 1999). If the model fits the data, we would expect the points to lie on the 45 degree line. We can see in Figure 7 that the points follow the 45 degree line very closely except for some large values of time, which is not a matter of concern for the censored data. Therefore, we can conclude that the graph shows an overall good of fit to the data.

Figure 7: Cox-Snell residual goodness of fit of the model



3.2.1 Diabetes

Diabetes status was not collected by ANZDATA until 1990. To check whether diabetes was also a risk factor, the data since 1990 were reanalyzed with diabetes status included as a variable in the model. The same model process as before was applied. Table 5 shows the results of this model, and it can be seen that the hazard ratios were almost the same as the model using the data from all years. The risk of dying of myeloma patients who were diabetic was 0.89 times that of the myeloma patients who did not have any diabetes, which was not statistically significant. Therefore, it was evident that there exhibited no significant effect of diabetes on survival of myeloma patients.

Table 5: Risk of death of myeloma patients since 1990

Covariates	**Hazard ratio**	**P-values**	**95% CI**
Age			
<60	1(referent)	-	-
60-70	1.22	0.139	(0.94, 1.60)
>=70	1.71	0.000	(1.32, 2.21)
Sex			
Female	1(referent)	-	-
Male	1.15	0.191	(0.93, 1.41)
Primary kidney disease			
Multiple myeloma	1(referent)	-	-
Possibly myeloma related	0.63	0.026	(0.42, 0.94)
GN/IgA	0.51	0.008	(0.31, 0.84)
Other	0.68	0.035	(0.47, 0.97)
Diabetes			
No	1(referent)	-	-
Yes	0.89	0.408	(0.67, 1.18)

3.3 Estimating the risk of dying of myeloma vs. non-myeloma patients on dialysis

This section summarises the analysis that was conducted to estimate the risk of dying between myeloma compared to non-myeloma patients on dialysis. The reason for comparing myeloma and non-myeloma patients was to estimate the risk of dying between myeloma and non-myeloma patients with respect to dying once on dialysis, after adjusted for confounding factors.

For the analysis, time was measured from diagnosis of ESKD to death. Figure 8 show that myeloma is very uncommon under the age of 40. As a result, the model only included those who were 40 years or older so that the non-myeloma and myeloma patients were comparable with respect to age. Thus this model has been developed for myeloma and non-myeloma patients, who are at least 40 years, as the myeloma is rare under the age of 40 (Leukaemia foundation, 2006). Also, I have excluded patients of ESKD before 1990 as the starting year of diagnosis of diabetes so that myeloma and non-myeloma patients were comparable with respect to year.

Figure 8: Histogram of age of myeloma and non-myeloma patients

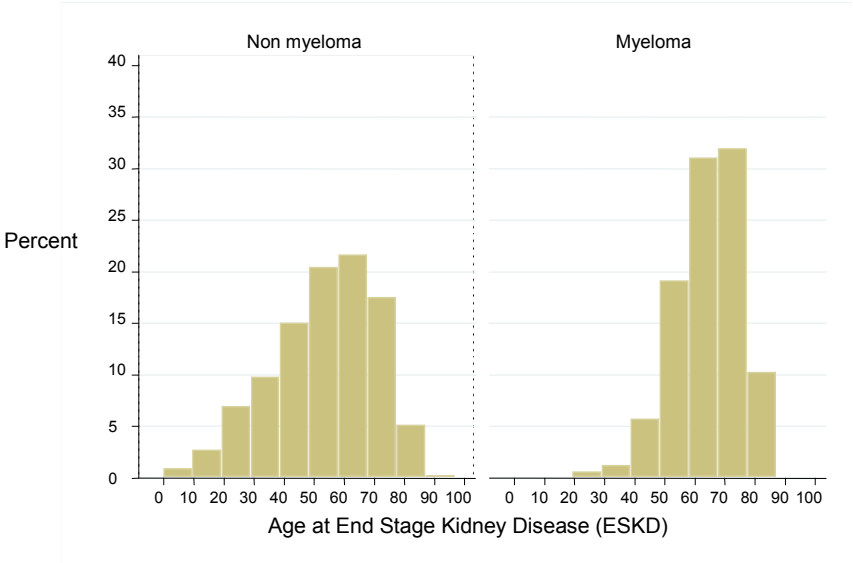
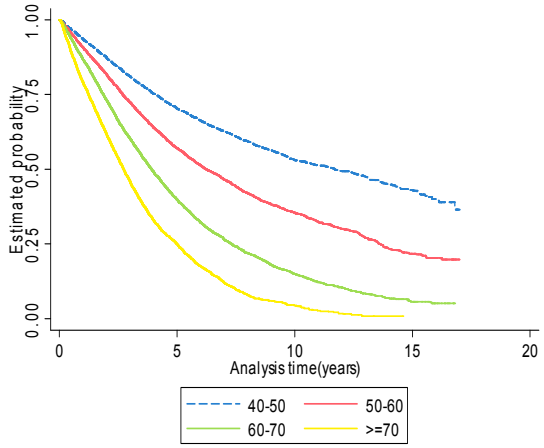
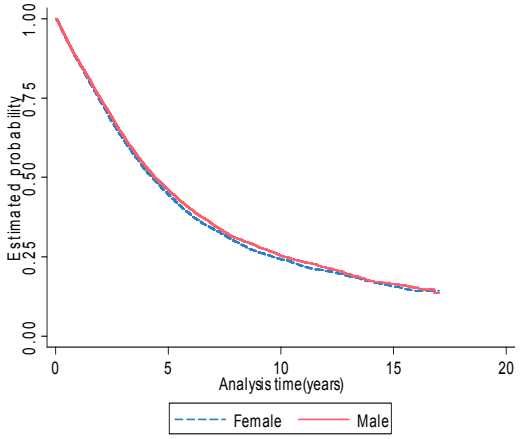


Figure 9 show that there is a difference in survival in terms of age group, types of dialysis, diabetes, smoking status, year of diagnosis of myeloma and myeloma vs. non-myeloma group, while there appears to be no or little difference in terms of gender, ethnicity, year of diagnosis of ESKD.

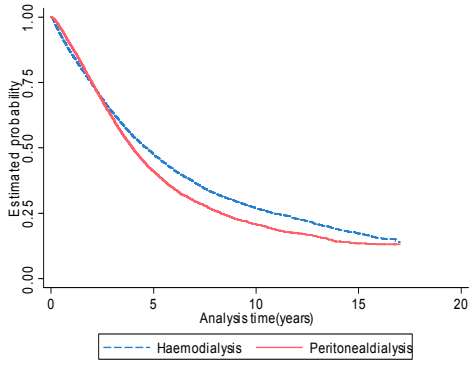
Figure 9: Kaplan-Meier Survival curves of the covariates of myeloma and non-myeloma patients



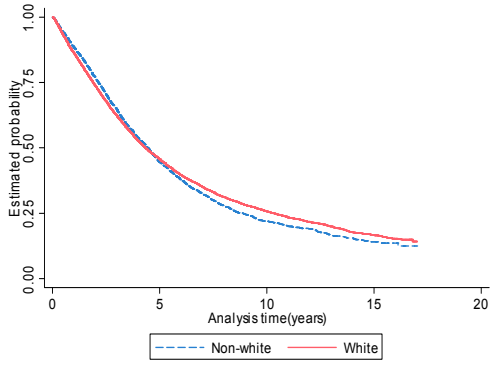
a) Probability of survival between different age group



b) Probability of survival between male and female

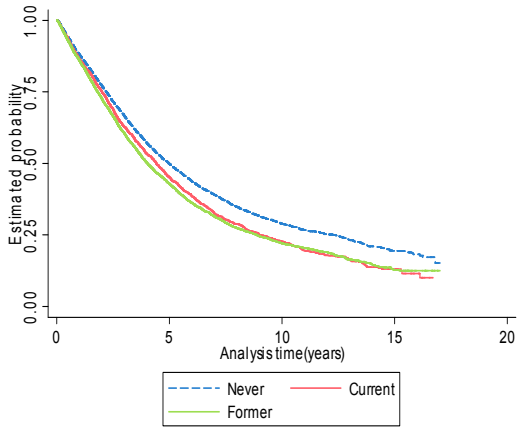


c) Probability of survival between haemodialysis and peritoneal dialysis

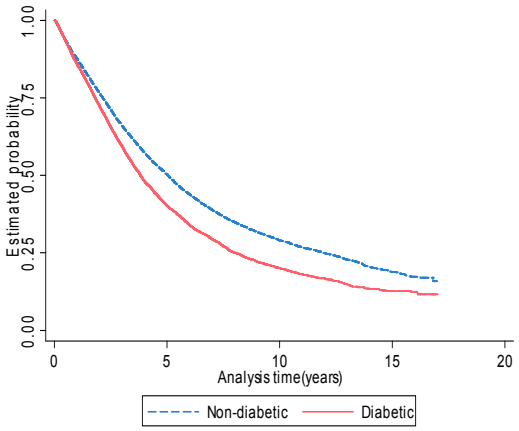


d) Probability of survival between white and non-white patients.

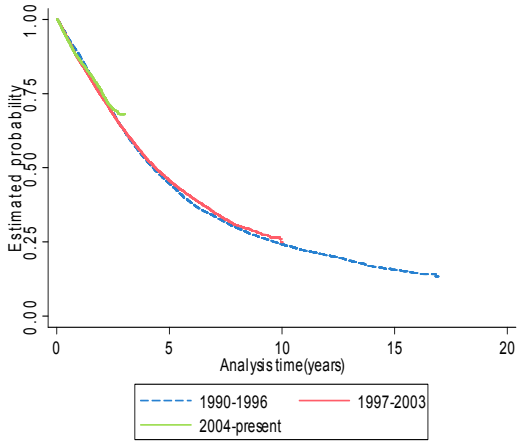
Figure 9 continued



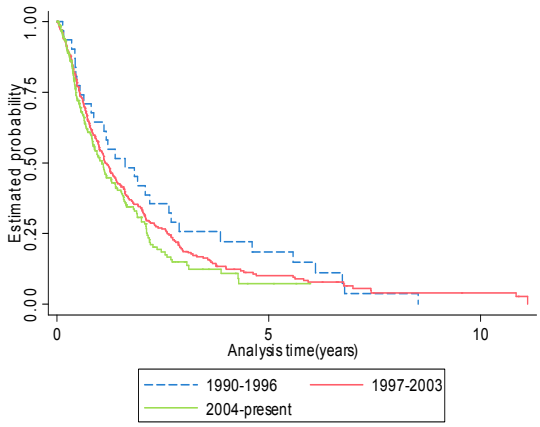
e) Probability of survival between different smoking status patients



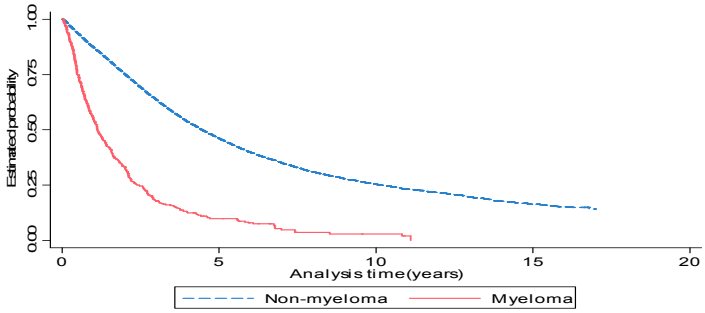
f) Probability of survival between diabetic and non-diabetic



g) Probability of survival between year of diagnosis of ESKD



h) Probability of survival between year of diagnosis of Myeloma



i) Probability of survival between myeloma and non-myeloma patients

After developing K-M curves and log-rank test of the covariates, we performed univariate and multivariate analyses respectively where covariates that had a $p<0.25$ went in to the multivariate model and those with $p<0.05$ stayed in the multivariate model.

The results of the univariate and multivariate analysis are given in Table 7.

Table 7: Tests of significance of Univariate and Multivariate analysis for myeloma and non-myeloma patients using Wald tests

Variables	***DF***	***P<=0.25 (Univariate)***	***P<=0.05 (Multivariate)***
Age group	3	<0.001	<0.001
40-50			
50-60			
60-70			
>70			
Gender	1	0.03	<0.001
Female (referent)			
Male			
Smoking status	1	<0.001	<0.001
Never(referent)			
Current			
Former			
Diabetes mellitus	1	<0.001	<0.001
No(referent)			
Yes			
Treatment type	1	<0.001	<0.001
Haemodialysis (referent)			
Peritoneal dialysis			
Myeloma	1	<0.001	<0.001
Year of diagnosis of ESKD	2	0.11	
1990-1996 (referent)			
1997-2003			
2004-present			

After developing the main effects model, we included interactions terms if $p<0.05$. We specifically tested only for interactions that seemed plausible with input from Dr Webster, which were myeloma with age and sex.

Comparison between model with interaction and model without interaction:

The likelihood-ratio test was used to compare the models in terms of fitting data. Clearly, it is seen that the interaction of myeloma with sex in model B is not significant (p=0.11). Therefore, it will be more logical, if we compare the models that are significant. Accordingly, table 8 shows that the Akaike's Information Criterion (AIC) of model A, which included sex, smoking status, treatment type, diabetes and the interaction of myeloma with age, is less than that of the model C, the main effect model, which included age, sex, smoking status, treatment type, diabetes and myeloma . Therefore, model A, the interaction model, is preferable to model C. The chi-square value of this test is 27.12 with p<0.001. That is, model A fits the data better than model C.

Table 8: Likelihood-Ratio Test of Models with and without the interaction terms

Model	Model type	Model Elements	P-values	Log-Log (null)	Log-Log (model)	DF	AIC
C	Main effect	Age+sex+smoking status+treatment type+diabetes+myeloma	<0.001	-132352.1	-130153.5	9	260324.9
B	Main effect+Interaction	Myeloma × sex	0.11	-132352.1	-130152.2	10	260324.4
A	Main effect+Interaction	Myeloma × age	<0.001	-132352.1	-130139.9	12	260303.8

Model diagnostics:

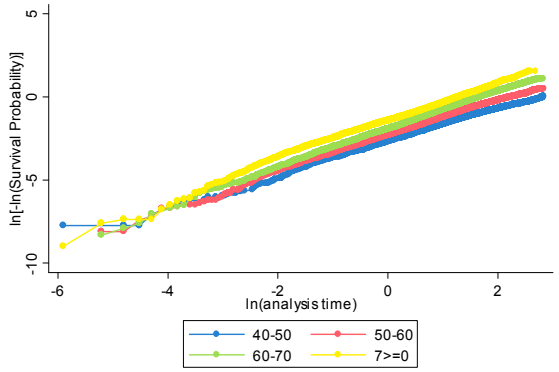
We performed diagnostic tests based on Schoenfeld for overall and scaled Schoenfeld residuals for each covariate to tests whether PH assumption is met. The output of the test is given below:

Table 9: Test of proportional hazards assumption

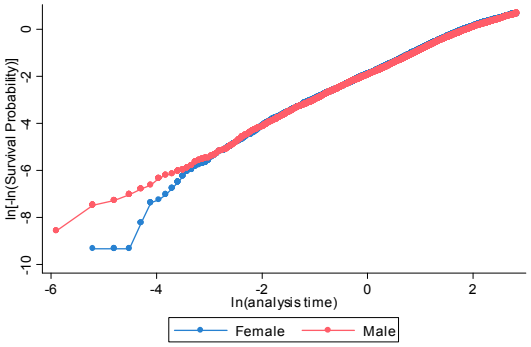
Variables	*Correlation coefficient (Rho)*	*Chi-squares*	*Degrees of freedom*	*P-values*
Age group				
40-50 (referent)				
50-60	0.032	14.27	1	<0.001
60-70	0.063	57.88	1	<0.001
>70	0.058	48.12	1	<0.001
Myeloma by age group				
40-50 (referent)	-0.008	0.95	1	0.330
50-60	0.004	0.27	1	0.606
60-70	-0.005	0.31	1	0.576
>70	0.005	0.32	1	0.571
Sex				
Female (referent)				
Male	0.021	6.34	1	0.012
Smoker				
Never(referent)				
Current	0.019	5.27	1	0.022
Former	-0.008	0.83	1	0.363
Diabetes				
No(referent)				
Yes	-0.004	0.19	1	0.660
Treatment type				
Haemodialysis (referent)				
Peritoneal dialysis	0.044	27.53	1	<0.001
Global Test		121.66	12	<0.001

It is seen that the global test is highly significant with $p<0.001$. Therefore, it is hard to say that the overall PH assumption is met. Again, all the covariates met PH assumption except *sex, treatment type and age group*, as their p-values<0.05. As a result, it is recommended that the model needs to be stratified by those covariates. However, we will look at the graphical representation, an alternative of the statistical test of PH assumption.

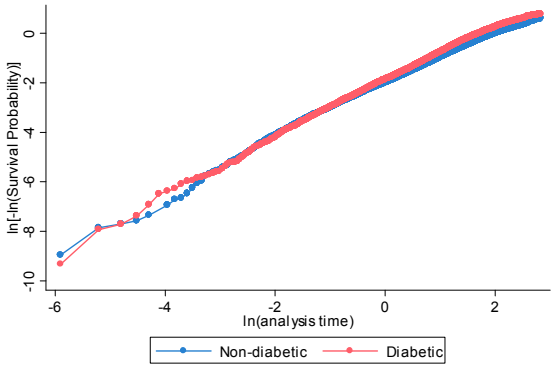
Figure 9: Log-Log survival curves for PH test



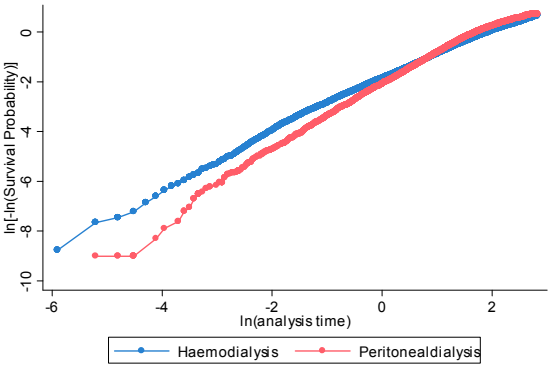
a) Age group



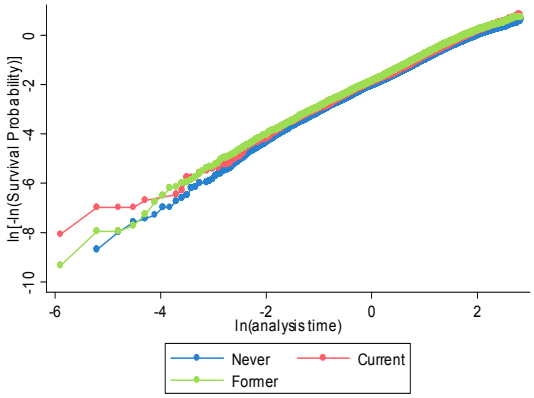
b) Gender



c) Diabetes



d) Treatment type



e) Smoker

Figure 9 shows that all the graphs are almost straight and parallel except age and treatment type, where the plot still indicates that PH may not hold as lines crosses. As a result, it is recommended that the model needs to be stratified by age and treatment type.

Stratified Cox Model:

The model is built up by stratifying age and treatment type respectively.

I. Cox Model Stratified by Age:

The output of the Cox model stratified by age group is given in Table 10.

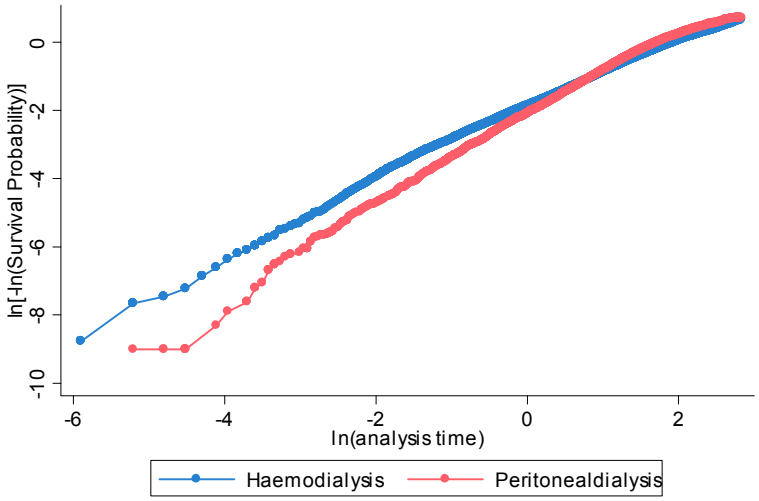
Table 10: Age adjusted hazard ratios of death of myeloma and non-myeloma patients

Variables	**Hazard ratio**	**P-values**	**95% CI**
Myeloma by age group		<0.001	
40-50	8.30		(5.66, 12.17)
50-60	4.57		(3.56, 5.86)
60-70	3.26		(2.72, 3.92)
>70	3.10		(2.66, 3.61)
Sex		<0.001	
Female (referent)			
Male	0.92		(0.89, 0.95)
Smoker		<0.001	
Never(referent)			
Current	1.53		(1.45, 1.62)
Former	1.19		(1.14, 1.23)
Diabetes		<0.001	
No(referent)			
Yes	1.38		(1.33, 1.43)
First treatment type		<0.001	
Haemodialysis (referent)			
Peritoneal dialysis	1.11		(1.07, 1.15)

Table 11: Test of proportional hazards assumption after stratifying by Age

Variables	*Correlation coefficient (Rho)*	*Chi-squares*	*Degrees of freedom*	*P-values*
Treatment type				
Haemodialysis (referent)				
Peritoneal dialysis	0.045	28.69	1	<0.001

Figure 10: Log-Log survival curves of treatment type for PH test after stratifying by Age



From table 11 and figure 10, it is seen that the treatment type still does not satisfy the proportional hazard assumption even after stratifying by age group. Therefore, the model will be stratified by age and treatment type.

II. Cox Model stratified by age and treatment type:

The output of the Cox model stratified by age and treatment type is given in Table 12.

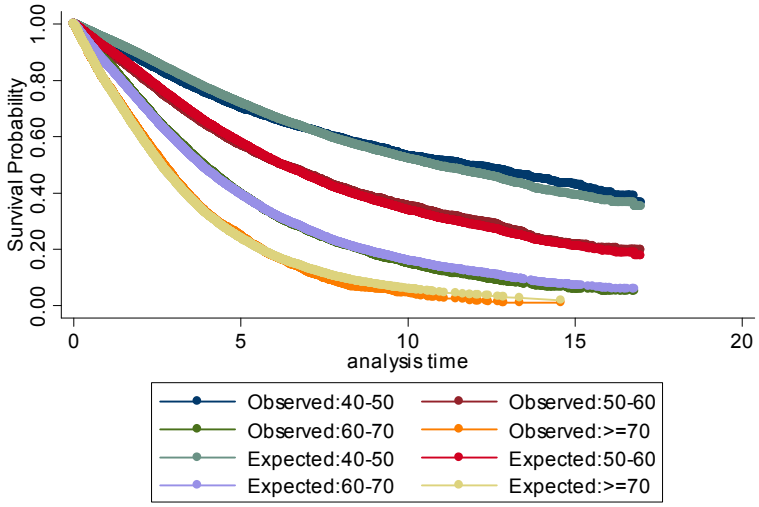
Table 12: Age adjusted hazard ratios of death of myeloma and non-myeloma patients

Variables	Hazard ratio	P-values	95% CI
Myeloma by age group		<0.001	
40-50	8.13		(5.54, 11.93)
50-60	4.54		(3.54, 5.83)
60-70	3.19		(2.66, 3.84)
>70	3.03		(2.60, 3.53)
Sex		<0.001	
Female (referent)			
Male	0.92		(0.89, 0.95)
Smoker		<0.001	
Never(referent)			
Current	1.53		(1.45, 1.62)
Former	1.19		(1.15, 1.23)
Diabetes		<0.001	
No(referent)			
Yes	1.38		(1.33, 1.43)

From table 10 and table 12, it is seen that the hazard ratios of death of myeloma and non-myeloma patients are almost similar. Therefore, the model stratified by age and the model stratified by age and treatment type provide the similar results. As a result, the final model can be built up either stratifying by age group only or it can be built up stratifying by both age and treatment type. However, if the study needs to estimate the risk of dying between myeloma and non-myeloma patients in terms of treatment type, the model needs to be stratified by age group only, because the model stratified by treatment type does not provide the estimate of treatment type.

Figure 11 shows the justification of the Cox model stratified by age:

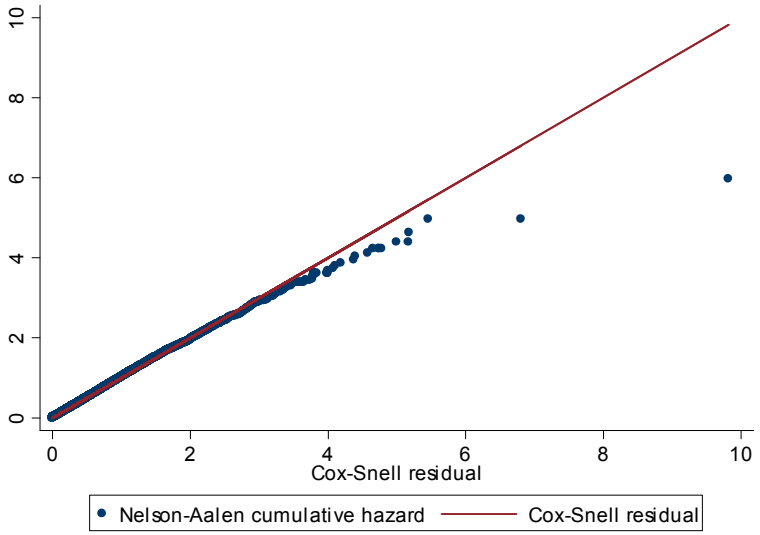
Figure 11: Comparison between Kaplan-Meier survival curves and Cox adjusted survival curves



We can see that the observed and expected lines are almost matching with each other for the respected age group. Therefore, the stratified Cox model is justified using age as a stratification variable.

Figure 12: Overall goodness of fit of the final model

To test whether the model fits the data well, we performed Cox-Snell residual goodness fit test, which can be seen as below. If the model fits the data, we would expect the points to lie on the 45 degrees line.



We can see that the points follow the 45 degrees line very closely except for some large values of time, which is not a matter of concern for the censored data. Therefore, we can conclude that the graph shows an overall good of fit to the data.

Interpretation of results:

The hazard ratio of age groups 40-50, 50-60, 60-70 and >70 are 8.30, 4.57, 3.26 and 3.10 respectively. That is the risk of dying in the myeloma compared to non myeloma group decreased with increasing age. Specifically, a patient aged 40-50 years with myeloma is 8.30 times more likely to die than a similar patient without myeloma. A patient aged 50-60 years with myeloma is 4.57 times more likely to die than a similar patient without myeloma. A patient aged 60-70 years with myeloma is 3.26 times more likely to die than a similar patient without myeloma. For patients aged over 70 years, the risk of dying in myeloma patients is 3.10 times that of patients without myeloma. The risk of dying is 8% lower in male than that of the female. The risk of dying of the patients with myeloma who are currently smoking is 1.53 times that of the patients without myeloma who never smoked. Again, the risk of dying of the patients with myeloma, who were smoking formerly, is 1.19 times that of the patients without myeloma who never smoked. The risk of dying of the patients with myeloma who had diabetes is 1.38 times that of the patients without myeloma who did not have diabetes. The result also shows that the risk of dying of the patients with myeloma who were treated with peritoneal dialysis is 1.11 times that of the patients without myeloma who were treated with haemodialysis. It can be mentioned here that all the results were highly significant ($p<0.001$).

4. Discussion

The results of the analyses presented in the report were consistent with the findings of other studies. For example, Leukaemia foundation found that myeloma was more common in men than in women (Leukaemia Foundation, 2006). They also found that myeloma was uncommon under the age of 60 years. In the ANZDATA registry, we found that myeloma was 62.2% common in men and 37.8% in women. We also found that majority (68.4%) of people diagnosed with myeloma were over 60 years of age.

This report interpreted the results of hazard ratio based on two different models. Model 1 (Section 3.2) analysed the myeloma patients only and model 2 (Section 3.3) compared the myeloma patients with non-myeloma patients on dialysis. In terms of the first analysis (myeloma only), only three risk factors were found to be statistically significantly associated with the survival of myeloma patients: age, primary renal disease and time of diagnosis of myeloma. The risk of dying was higher if the primary kidney disease caused by myeloma compared to other diseases.

The second analysis focused on comparing the mortality of myeloma with non-myeloma patients. Specifically, the hazard ratios for myeloma patients were decreasing with increasing age when compared with non-myeloma patients, whereas the hazard ratios were increasing with increasing age when only the myeloma patients were analysed. Therefore, the risk of dying for myeloma patients is increasing with age in absolute terms, but in relative terms the risk of dying, although always higher compared to non-myeloma patients, becomes smaller with age.

In terms of the analysis which compared myeloma with non-myeloma dialysis patients, there were potentially other approaches that could have been taken, besides restricting the age to 40 years or older. One option was to analyse the data for all ages. However, we were concerned that the results from this approach may have been biased, as myeloma does not occur in children and rarely under the age of 40. Another approach might have been to use relative survival. Relative survival is the ratio of the observed survival of all causes in the patient group and the expected survival of a comparable group from the general population (Simpson and McGeehan, 2008). Usually, the general population is the national population, for example, the Australian population. However, we wished to compare myeloma with the non-myeloma dialysis patients. Hence, in our case the general population would have been the entire dialysis population. This approach is worth further investigation but is beyond the scope of this project.

References:

ANZDATA. (2004). Available at http://www.anzdata.org.au/ANZDATA/anzdatawelcome.htm Last accessed on 09 Aug 2004.

Claire et al. (2006). Cancer Incidence Before and After Kidney Transplantation. The Journal of the American Medical Association, Vol. 296 (23), pp.2823-2831. Available at http://jama.ama-assn.org/cgi/content/full/296/23/2823.

Hosmer DW & Lemeshow S. (1999). Applied Survival Analysis: Regression modelling of time to event data. New York: Wiley.

Harrell F. E. (2006). Regression Modelling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. USA: Springer.

Kleinbaum DG & Klein M. (1996). Survival Analysis: A self-learning text. USA: Springer.

Leukaemia Foundation. (2006). Understanding Myeloma: A guide for patients and families. Available at http://www.leukaemia.org.au/web/resources/pdf/booklet_myeloma.pdf. Last accessed on May 2006.

Nephrology Dialysis Transplantation. (1996). Multiple myeloma treatment in dialysis-dependent patients: to transplant or not to transplant? Oxford Journal, Vol. 11(7), pp. 1486-1487. Available at http://ndt.oxfordjournals.org/cgi/reprint/11/7/1486.

Simpson JM & McGeehan K. (2008). Survival Analysis Notes. University of Sydney: School of Public Health.

University of Maryland Medical Centre. (2007) End-stage kidney disease – Overview. Available at http://www.umm.edu/ency/article/000500.htm. Last accessed on 14 Aug 2004.

Appendix

Table 13: Description of the variables and definitions received from ANZDATA:

Variables	*Description*
patient	ANZDATA patient ID
disease	Primary renal disease
frsttreatdate	Date 1st Renal Replacement Therapy (RRT) treatment
catype	Histological cancer type
cadate	Date cancer diagnosed
agecancer	Age cancer diagnosed
castage	Stage at diagnosis
sex	Gender
race	Ethnicity; white, non-white
raceother	Race others
firsttxdate	First transplant date
Firstdonor	First donor of kidney
tyedescr	Type of cancer description
causeprd	Cause of primary renal disease
candeath	Date of death of cancer
causedeath	Cause of death of cancer
smoker	Smoker at 1st RRT
diabetes	Diabetes Mellitus
diabdate	Date of diabetes

initstat	State resident at 1st RRT
statedia	State resident at cancer diagnosis
primsite	Cancer site code
last	Last known outcome
datdeath	Date of last known outcome
initstat	State resident at 1st RRT
statedia	State resident at cancer diagnosis
status	Death or alive
era	Year of 1st transplant
fuptime	Duration from Tx to last known
prd	Primary cause of ESRF
Agetxcat	Age categorisation around age at tx quartiles, for those with cancer 'events'
rrttime	Duration RRT prior to 1st tx
cancer	cancer ever, no CIN (cervical cancer)
firsttreat	1st treatment code
sequ	Patient sequence
sequtot	Patient max sequence

Table 14: Description of the variables and definitions derived and defined after receiving ANZDATA:

New/Re-categorisation of variables	*Description*
firsttreat_type	Categorisation of 1st treatment separating dialysis and transplant patients
firstdxdate	First treatment date dialysis only
myeloma	Separating myeloma cancer using cancer type=10
mycadate	Myeloma cancer date
myeloma_ever	Separating myeloma and non-myeloma patients
timebetn	Generating time between getting myeloma & other cancer
eskd_between	Time between myeloma and End Stage Kidney Disease(ESKD)
fuptime	Duration from dialysis to last known
status	Dead(1) and alive or temp recovery or lost(0)
prd	Categorisation of primary cause of ESKD
era	Year of 1st dialysis
agedx	Age at 1st dialysis
agedxcat	Categorisation of age at 1st dialysis
myelomaeskd	Duration of Myeloma diagnosis
era_eskd	Year of ESKD date
era_eskdgp	Categorisation of era_eskd
era_myeloma	Year of Myeloma cancer date
era_myelomagp	Categorisation of era_myeloma
pkd	Re-categorisation of prd into smaller group