UNIVERSITY OF SOUTHAMPTON
FACULTY OF SOCIAL AND HUMAN SCIENCES
Mathematics

Informative Censoring in Transplantation Statistics

by

Natalie Staplin

Thesis for the degree of Doctor of Philosophy

March 2012

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL AND HUMAN SCIENCES

Mathematics

Doctor of Philosophy

INFORMATIVE CENSORING IN TRANSPLANTATION STATISTICS

by Natalie Dawn Staplin

Observations are informatively censored when there is dependence between the time to the event of interest and time to censoring. When considering the time to death of patients on the waiting list for a transplant, particularly a liver transplant, patients that are removed for transplantation are potentially informatively censored, as generally the most ill patients are transplanted. If this censoring is assumed to be non-informative then any inferences may be misleading.

The existing methods in the literature that account for informative censoring are applied to data to assess their suitability for the liver transplantation setting. As the amount of dependence between the time to failure and time to censoring variables cannot be identified from the observed data, estimators that give bounds on the marginal survival function for a given range of dependence values are considered. However, the bounds are too wide to be of use in practice. Sensitivity analyses are also reviewed as these allow us to assess how inferences are affected by assuming differing amounts of dependence and whether methods that account for informative censoring are necessary. Of the other methods considered IPCW estimators were found to be the most useful in practice.

Sensitivity analyses for parametric models are less computationally intensive than those for Cox models, although they are not suitable for all sets of data. Therefore, we develop a sensitivity analysis for piecewise exponential models that is still quick to apply. These models are flexible enough to be suitable for a wide range of baseline hazards. The sensitivity analysis suggests that for the liver transplantation setting the inferences about time to failure are sensitive to informative censoring. A simulation study is carried out that shows that the sensitivity analysis is accurate in many situations, although not when there is a large proportion of censoring in the data set.

Finally, a method to calculate the survival benefit of liver transplantation is adapted to make it more suitable for UK data. This method calculates the expected change in post-transplant mortality relative to waiting list mortality. It uses IPCW methods to account for the informative censoring encountered when estimating waiting list mortality to ensure the estimated survival benefit is as accurate as possible.

Contents

List of Figures

List of Tables

Author's Declaration

I, Natalie Dawn Staplin, declare that the thesis entitled 'Informative Censoring in Transplantation Statistics' and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- none of this work has been published before submission;

Signed:

Date:

Acknowledgements

I would like to thank my supervisors, Alan Kimber, Dave Collett and Paul Roderick for all their guidance during my PhD.

I would like to thank NHS Blood and Transplant for providing the funding for my thesis and also supplying all the data required for the project.

Finally, I would like to thank my parents, Stewart and Lorraine Staplin, for all their support, both financial and emotional, throughout my studies. It is to them that I dedicate this thesis as I could not have done it without them.

List of Acronyms

AID Auto-immune and cryptogenic disease

ALD Alcoholic liver disease

ANOVA Analysis of variance

CAR Coarsened at random

CI Confidence Interval

CVA Cerebrovascular accident

DRI Donor risk index

ECD Extended criteria donors

ESLD End-stage liver disease

ESRF End-stage renal failure

HBV Hepatitis B cirrhosis

HCV Hepatitis C cirrhosis

IFALD Intestinal Failure Associated Liver Disease

INR International normalized ratio

IPCW Inverse probability of censoring weighted

ISNI Index of sensitivity to non-ignorability

KM Kaplan-Meier

LHS Left hand side

MELD Model for end-stage liver disease

MLE Maximum likelihood estimate

MPL Modified partial likelihood

NHSBT National Health Service Blood and Transplant

PBC Primary biliary cirrhosis

PLD Primary liver disease

PSC Primary sclerosing cholangitis

RHS Right hand side

RTA Road traffic accident

SE Standard error

TOD Time of death

UKELD United Kingdom model for end-stage liver disease

UKTR United Kingdom Transplant Registry

Chapter 1

Introduction

Survival analysis methods are often used in the analysis of medical data, both designed clinical trials and observational studies, as the event of interest may not be completely observed. This could be because individuals drop out of a study, experience a different event that means it is no longer possible to observe the event of interest, or the event of interest has not been observed by the end of the study. Any individual for whom the event of interest is not observed is censored at the time that they are removed from the study. Kalbfleisch and Prentice (2002) and Collett (2003) both give detailed introductions to the analysis of data with censored observations, some of which is summarised in this section.

The time from a given origin to the event of interest is represented by the random variable, T and the time from the same origin to censoring is represented by C. The time to event is often referred to as time to failure in the literature. A value of interest in most survival analysis applications is the marginal survival function of T, $S_T(t) = P(T > t)$. This could be either the overall survival function or the survival function for a particular individual. The factors that affect the survival time are also usually of interest, so where possible any methods used should allow for the inclusion of covariates. Using suitable models, the effect of any prognostic factors on the survival time can be estimated.

In this thesis, we shall consider the case where both T and C are continuous. Therefore, we will not be looking at Type I censoring, where the censoring time of each individual is fixed in advance.

Due to the censoring, a value of T for each individual may not be observed. This means some assumption about the association between T and C needs to be made before $S_T(t)$ or any model parameters can be estimated. The standard assumption used in survival analysis methods is that of non-informative censoring. This means that T and C are independently distributed, so as described in Kalbfleisch and Prentice (2002), the conditional hazard function satisfies

$$h_C(t|T, T > t) = h_C(t|T > t), \tag{1.1}$$

where

$$h_C(t) = \lim_{\delta t \to 0} \left\{ \frac{P(t \leq C < t + \delta t | C \geq t)}{\delta t} \right\}$$

is the hazard function for censored observations. This means censored observations provide only the information that the true survival time exceeds the survival time and no information about the subsequent survival time. This assumption is suitable for some types of censoring, as long as the censoring occurs randomly. For example, when individuals are censored because they are still at risk at the end of the study, it is reasonable to assume that this type of censoring is non-informative. Another example of non-informative censoring occurs when individuals who are removed from the study when they experience events unrelated to the event of interest, such as being run over by a bus.

However, there are other types of censoring where this assumption may be questioned and models that allow for dependence between T and C need to be used. There are many issues that are raised when fitting these types of models which will be discussed in the following section.

1.1 Informative Censoring

If it assumed that there is dependence between T and C and the conditional hazard function for C does not satisfy (1.1), then the censoring is called informative. One possible cause of informative censoring is that the factors that control time to censoring are also associated with time to event.

There are several situations that have been identified where censoring is likely to be informative. According to Lagakos (1979), these are

1. when individuals drop out of a clinical trial for reasons that could be related to the therapy,

2. when individuals are removed from a clinical trial by design and no longer followed for survival time if they experience a specific critical event, and

3. when individuals in a study experience a failure from a cause of secondary interest which censors the failure time from the cause of primary interest.

The type of drop-out in situation 1 can generally not be avoided even if measures are put in place to minimize the number of patients who leave a study before the end. In situation 2, the specific critical event is defined by those designing the study, such as the spread of disease past a given threshold. So, the difference between situations 2 and 3 is that the censoring event in situation 2 can be avoided by using different study designs whereas the

censoring event in situation 3 is unavoidable. Situation 3 here is known as the problem of competing risks which is considered in more detail in Section 1.1.1.

The standard methods used to analyse censored data are based on the assumption of non-informative censoring and may not be robust to the assumption of informative censoring. In the non-informative censoring case, we assume that those who are censored are representative of the individuals who are at risk at the time of censoring. If there is a positive association between time to event and time to censoring then those who are censored would have a smaller expected survival time. This could lead to the standard methods overestimating the survival function. Conversely, a negative correlation between time to event and time to censoring could lead to underestimation of the survival function when using standard methods. The robustness of the standard methods would also be affected by the proportion of observations that are informatively censored. It is found in Fisher and Kanarek (1974) using simulated data that the more informative censoring there is in a data set, the stronger the effect of informative censoring on the underlying survival function. This means that the standard methods would generally be more biased as the amount of informative censoring in a data set increases.

One possible way of incorporating informative censoring into a model is to use a bivariate distribution for (T,C) that has independence of T and C as a special case. However, it is not possible to test which bivariate distribution should be used in a particular application due to identifiability issues. The implications of these issues are discussed in Section 1.1.2.

1.1.1 Competing risks framework

It is possible to use a competing risks framework when we have censored data, this is described in Crowder (2001), which is where the definitions given in this section are taken from. In a competing risks framework, there are m possible event types that could be observed. Here, the situation where only one event type can be observed for each individual is considered. Therefore, for the ith individual in the data set, the observed data are the event time Y_i and the event type J.

The overall hazard function at time t is

$$h(t) = \lim_{\delta t \to 0} \left\{ \frac{P[t \leq T < t + \delta t | T \geq t]}{\delta t} \right\}.$$

To model the competing risks, the sub-hazard function

$$h(j,t) = \lim_{\delta t \to 0} \left\{ \frac{P[t \leq T < t + \delta t, J = j | T \geq t]}{\delta t} \right\}, \quad j = 1, \ldots, m,$$

is used. This is the hazard rate for event type j at time t, in the presence of all the other

types of event. The sub-hazard functions are related to the overall hazard function as

$$h(t) = \sum_{j=1}^{m} h(j,t).$$

The literature that uses this framework often defines latent event times for each type of event, denoted $\bar{Y}_1, \ldots, \bar{Y}_m$. The actual observation time is given by

$$Y = \min(\bar{Y}_1, \ldots, \bar{Y}_m),$$

and the corresponding event type is J so that $Y = \bar{Y}_J$.

It is often the marginal distribution of one of the latent event times that is of interest. Generally, it is the marginal hazard functions

$$h_j(t) = \lim_{\delta t \to 0} \left\{ \frac{P[t \leq \bar{Y}_j < t + \delta t | \bar{Y}_j \geq t]}{\delta t} \right\}, \quad j = 1, \ldots, m,$$

that are considered, which are the hazard rates for event type j at time t in the absence of any other event types. These marginal hazard functions cannot be identified from the observed data.

The usual strong assumption that is made is that $\bar{Y}_1, \ldots, \bar{Y}_m$ are statistically independent. This means that the marginal hazard functions now equal the sub-hazard functions and are therefore identifiable. But the assumption used to achieve this is untestable.

The situation that is being considered here can be set up using the competing risks framework described here. In this case, there are only two competing events, with latent event times $\bar{Y}_1 = T$ and $\bar{Y}_2 = C$. The event time that is of interest is T, and the marginal distribution of this variable is what needs to be estimated. The observed data will be $Y = \min(T, C)$ and J which denotes the event type. However, as there are only two types of event, it usual that an indicator variable $\Delta = I(Y = T)$ is recorded instead of J.

1.1.2 Identifiability issues

It has already been mentioned in Section 1.1.1 that the marginal distributions of the latent event times of m competing risks are not identifiable. This is because there is insufficient information in the observed data to be able to identify the joint distribution of the latent event times $\bar{Y}_1, \ldots, \bar{Y}_m$. This was first discussed in Cox (1959), who considered the case with just two random variables that follow a general independent risks model. The random variables $\bar{Y}_1$ and $\bar{Y}_2$ are independently distributed with continuous distribution functions $F_1(t)$ and $F_2(t)$. Cox (1959) stated that "no data of the present type can be inconsistent with [the general independent risks] model".

Tsiatis (1975) considers the general case where there are m competing risks acting on the system. It is shown that for any given joint survivor function where there is

dependence between the variables, a different joint survivor function can be identified where the variables are independent. Both of these survivor functions give the same observable functions so it is not possible to distinguish between them.

Crowder (1991) extends the result above to show that each independent risks model has a class of satellite dependent models with the same observable functions. It is also shown that this class can be further broken down into sets with the same marginal functions. If it were possible to have unlimited observation of $Y = \min(T, C)$ and the corresponding indicator along with unlimited observation of T and C then we would be able to identify the subclass. If this subclass did not contain the independent risks models then at least it can be identified that there are not independent risks. However, for the medical examples of interest, this unlimited observation is not feasible.

Therefore, in general it is not possible to construct a statistical test for non-informative censoring with the alternative being informative censoring for the situation that is being considered here.

These problems of non-identifiability also have implications for any informative censoring models fitted. One of the most popular approaches is to specify a bivariate distribution for (T, C) for which independence of T and C is a special case. The parameters will no longer be unidentifiable, as long as each point in the parameter space of this joint distribution has a corresponding distinct distribution for the observed data (Y, Δ). However, due to the lack of information in the observed data, any joint distribution assumed for (T, C) cannot be verified using a statistical test.

1.2 Liver Transplantation

Much of the methodology developed in this thesis will be illustrated using data on liver transplantation candidates and recipients. Accordingly, in this section, an outline of relevant aspects of liver transplantation is given.

Liver transplantation is used as treatment for patients with end-stage liver disease (ESLD). In the UK, to receive a liver transplant, a candidate must normally be registered with NHS Blood and Transplant (NHSBT) and meet certain criteria as set out in Section 1.2.2. NHSBT are also responsible for the allocation of donor organs which is done according to the policy described in Section 1.2.3.

There has been an increase in the number of patients waiting for a liver transplant despite measures introduced with the intention of making a larger number of donor organs available. These include the use of extended criteria donors (ECD) that would have previously been rejected and split livers, so that one organ can be used to provide transplants for both an adult patient and a paediatric patient. Because of the short fall between the

number of donor organs and the number of patients on the waiting list for a transplant, it is important that any allocation policy should aim to maximise the utility of the organ whilst reducing the mortality of those on the waiting list. However, it is stated in Neuberger et al. (2008) that this may not be possible because "those who are very sick and thus at greatest risk of death may have a worse outcome and will use more resources".

There has been much discussion lately of the most appropriate method of allocating donor organs to transplant candidates. Freeman et al. (2009) compare the current centre-based policy UK policy, which is described in Section 1.2.3, and the patient-based US policy for allocating donor liver grafts. Neuberger et al. (2008) discuss the current UK policy but do say that in the long-term the aim is "to develop a model of allocation based on the greatest transplant benefit which would take into account both the likelihood of dying without a transplant as well as the likelihood of dying following a transplant". An allocation policy that is based on the idea of the greatest survival benefit from transplantation is discussed in Schaubel (2009a). There will be more discussion of survival benefit and an associated allocation policy in Section 1.2.6. There has also been a call for more transparency in the allocation policy in the UK by Elisabeth Buggins, Chair of the former Organ Donation Taskforce for Department of Health, in an independent report to clarify the rules on organ transplants for both NHS patients and non-UK EU residents.

If a policy that is based on statistical models is implemented, such as the allocation policy based on survival benefit, then the models of waiting list and post-transplant mortality used will need to be as accurate as possible. This is why methods that facilitate the implementation of good models are particularly useful, and this is part of the motivation behind the research carried out here.

1.2.1 United Kingdom model for End-Stage Liver Disease

A model to predict the severity of a patient's ESLD for the US was developed in Weisner et al. (2001), known as the model for end-stage liver disease or MELD. It uses three measurements: serum creatinine at time of registration, serum bilirubin at time of registration and the international normalized ratio (INR), which gives results of blood clotting tests. This model gives a score that reflects the measured level of liver dysfunction using the formula

$$\begin{aligned}\text{MELD} =& 9.57\log(\text{creatinine}(mg/dL)) + 3.78\log(\text{bilirubin}(mg/dL)) \\ &+ 11.2\log(\text{INR}) + 6.43. \end{aligned} \tag{1.2}$$

This score has also been found to be a significant predictor of mortality on the waiting list for a liver transplant.

A similar UK model for end-stage liver disease (UKELD) is described in Barber et al. (2007). Using an analysis of 1103 patients, the formula

$$\begin{aligned}\text{UKELD} =& 5\{1.5\log(\text{INR}) + 0.3\log(\text{creatinine}(\mu\text{mol}/L)) \\ &+ 0.6\log(\text{bilirubin}(\mu\text{mol}/L)) - 13\log(\text{Na}(mmol/L)) + 70\}\end{aligned} \tag{1.3}$$

was developed. It uses the same components as the MELD score as well as an additional component, serum sodium at time of registration. A UKELD score of greater than 49, calculated using (1.3), predicts a greater than 9% 1-year mortality. As a patient's UKELD score increases, their 1-year mortality will also increase. Patients with a UKELD score below 49 have a 1-year mortality of less than 9%.

The formula for the UKELD score has since been revised as is now given by

$$\begin{aligned}\text{UKELD} =& 5.395\log(\text{INR}) + 1.485\log(\text{creatinine}(\mu\text{mol}/L)) \\ &+ 3.130\log(\text{bilirubin}(\mu\text{mol}/L)) - 81.565\log(\text{Na}(mmol/L)) + 435\end{aligned} \tag{1.4}$$

as detailed in Barber et al. (2011). However, we use (1.3) to calculate the UKELD scores used in all the analyses in this thesis as the updated formula was not available at the time that we carried out the analyses. To assess whether our results are likely to be greatly affected by the use of the original UKELD score, in Section 1.2.5 we compare the values given by (1.3) and (1.4) for the data set used in this thesis.

1.2.2 Selection criteria for transplant waiting list

Patients who require a liver transplant are either registered for a super-urgent transplant or an elective transplant. The criteria for registration as a super-urgent transplantation are detailed in the Protocols and Guidelines for Adults Undergoing Deceased Donor Liver Transplantation in the UK, which is available on the NHSBT website (*http://www.nhsbt.nhs.uk/index.asp*). These are not considered here as these patients will not be included in any statistical models as they will always remain the top priority for any donor organ that becomes available.

Also detailed in the Protocols and Guidelines for Adults Undergoing Deceased Donor Liver Transplantation in the UK are the criteria for patients to be put on the waiting list for an elective transplant. To be accepted for an elective liver transplant, the candidate must have a projected 5-year survival after transplantation of at least 50%. Also, adult patients awaiting a first liver transplant must meet at least one of the following four criteria:

- Chronic liver disease or failure (UKELD score of 49 or greater)
- Hepatocellular carcinoma

- A variant syndrome
- Have been accepted through the National Appeals Panel

There are additional criteria for patients to be registered on the waiting list for an elective transplant other than a UKELD score of 49 or greater as the UKELD score does not always reflect the need for a liver transplant. For patients with hepatocellular carcinoma, the severity of this disease is not measured by the UKELD score so they need to be considered separately. The UKELD score does not incorporate quality of life, so any patients that need to receive a transplant to improve their quality of life must be considered separately.

1.2.3 Current allocation policy for donor organs

The current UK allocation policy for donor livers is summarised in Figure 1.1. The flowchart shows the order in which patients are considered when allocating a donor liver.

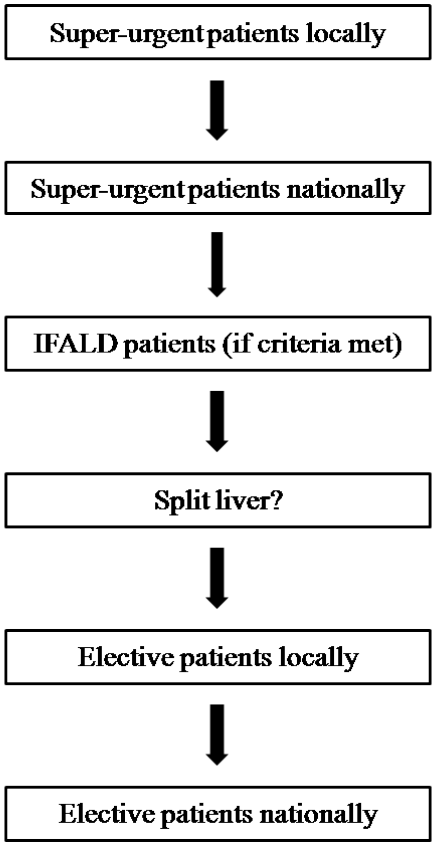

Figure 1.1: Flowchart showing the order of priority when allocating donor livers in the UK. (IFALD = Intestinal Failure Associated Liver Disease)

Patients that are registered on the list for a super-urgent transplant are given the highest priority. Organs are offered to super-urgent patients in the local area first before

being offered to super-urgent patients nationally. These patients are prioritised by the time spent on the super-urgent list. So, if there are several patients on the list, the organ will be given to the blood group compatible patient who has waited the longest.

If there are no suitable super-urgent patients or patients with intestinal failure associated liver disease on the list, then a decision needs to be made whether to split the liver to maximise the use of the donor organ. Generally, a donor liver is considered for splitting if the donor

- is less than 40 years old,
- weighs at least 50 kg and
- has been in intensive care for less than 5 days.

Once a decision has been made about splitting the liver, the organ will then pass to a unit, usually starting with the unit that covers the area where the organ is being retrieved. It is then the choice of the clinicians at this unit if there are any suitable patients registered at that unit and which of these patients the organ should be offered to. If there are no suitable patients at that unit, then the organ will be offered to another unit and so on in an agreed sequence until a suitable patient is found.

1.2.4 Issues arising when modelling survival of candidates on the waiting list for a liver transplant

The survival of adult patients on the waiting list for an elective liver transplant is of interest here. For obvious reasons, death on the waiting list will only be observed for a small subset of patients as the majority will be removed for transplantation and those with a deteriorating condition are likely to be removed from the list before death.

Therefore, a large amount of censoring will be observed in this situation. Those who are removed to receive a liver transplant are potentially informatively censored. The same can be said of those removed due to deteriorating condition. This is reasonable as patients who are removed for these reasons are generally the most sick on the list. They would have a higher risk of death and therefore a lower expected survival time than those who remain on the waiting list. This means that the estimated survival function may lie above the true survival function if such censoring is assumed to non-informative. Therefore the estimated probability of survival time at time t may be larger than the true survival probability at time t, or the the estimated survival function may overestimate the true survival function.

However, if a patient is removed from the list during the study for other reasons or because they were still active on the waiting list at the end of the study, then it will be assumed that this censoring is non-informative. This is a reasonable assumption for these patients as the censoring process is acting randomly here.

1.2.5 Liver Registration data set

A data set has been provided by NHSBT to illustrate the methods developed in this thesis. It is important to note that this data set is being used to motivate and illustrate the statistical methods and the results quoted should not be regarded as being definitive for guiding clinical practice. The data set contains data on all adult patients who were registered for an elective liver transplant between 1 January 2000 and 31 December 2008, which is 4594 rows of data. There were 203 patients that had multiple lines of data as they were re-registered for a liver transplant during this period. For example, they could have received a donor organ which later failed and were then put back on the waiting list for a transplant. The two registrations are then treated as two separate observations. This information was taken from the UK Transplant Registry (UKTR) on 7 April 2009. The date of registration on the waiting list is given along with the date of removal from the waiting list, and details of whether this removal was due to death, transplantation or for other reasons.

Some preliminary investigations of the data found that 39 patients were recorded as have been transplanted a few days before their time of registration or on the date of their registration. It was found that these patients received their transplant before their registration details were sent to NHSBT and entered into the UKTR. Therefore, these patients are removed from the data used here, as their time from registration to time of transplant would be recorded as non-positive. There were also two patients who were recorded as being registered on the list and being removed from the list on the same date, as well as three patients who were recorded as being registered on the list on the same day that they died. These five patients were also removed from the data set as they would have had an observation time of zero. This means that there were 4550 rows of data used in the analyses here.

Many covariates are also included in the data set, the details of which are given in Tables 1.1 to 1.9. Some of the earlier registrations in this data set do not have all the covariates recorded as the data collected at time of registration changed over the period under consideration.

Variable Name	Description and Details
RECIP_ID	Unique patient number
UNIT	Anonymised transplant centre (Levels: A (n=260), B (n=725), C (n=586), D (n=410), E (n=1109), F (n=1033), G (n=417))
ENDSTAT	Code denoting the current or final state of a registration (Levels: A Active (n=165), S Suspended (n=2), T Transplanted (n=3498), R Removed (n=451), D Died (n=478))
ADATE_ON	Date of first active record per registration
SDATE_ON	Date of first suspended record per registration
R_DATE	Date of removal from the transplant list
TX_DATE	Date of liver transplant
D_DATE	Date of death
RUN_DATE	Date registration dataset run (7 April 2009)
REG_AGE	Age at time of registration (Range: 17-78 years, Mean: 51.8 years)
LIV_DIS	Primary liver disease at time of registration (See Table 1.3)
PD_CAT	Primary liver disease grouped at time of registration (See Table 1.4)
PRIM_COD	Primary cause of death (See Table 1.5)
RCOD_GRP	Primary cause of death grouped (See Table 1.6)
RHEIGHT	Height at time of registration (Range: 62-205 cm, Mean: 169.75 cm, 208 missing)
RWEIGHT	Weight at time of registration (Range: 30-178 kg, Mean: 76.44 kg, 80 missing)
RSEX	Sex (Levels: 1 Male (n=2881), 2 Female (n=1713))
RBG	Blood group (Levels: 0 O (n=2053), 1 A (n=1831), 2 B (n=532), 3 AB (n=178))
RETHNIC	Ethnicity (Levels: 1 White (n=4077), 2 Asian or Asian-British (n=332), 3 Black or Black-British (n=103), 4 Chinese/Oriental (n=34), 7 Other (n=48))
LIVER	Completeness of liver transplanted (Levels: 0 Whole (n=3239), 1 Reduced (n=6), 2 Split (n=253), 1096 missing)
CREAT_REG	Serum creatinine at time of registration (Range: 7.6-400 μmol/l, Mean: 98.1 μmol/l, 1875 missing)
INR_REG	INR at time of registration (Range: 1-12, Mean: 1.5, 1902 missing)
BILIRUBIN_REG	Serum bilirubin at time of registration (Range: 1-1270 μmol/l, Mean: 98.7 μmol/l, 1871 missing)
SODIUM_REG	Serum sodium at time of registration (Range: 105-150 mmol/l, Mean: 135.9 mmol/l, 1887 missing)
UKELD_REG	UKELD score at time of registration (Range: 38-83.2, Mean: 55.5, 1924 missing)

Table 1.1: The variables applicable to all patients in the Liver Registration data set and giving details about the variables. For continuous variables the range and mean are given and for factorial variables the levels of the factor and the number of observations at each level are given. The number of observations missing the covariate value is also noted.

Variable Name	Description and Details
COF	Cause of graft failure (See Table 1.7)
FAILDATE	Date of graft failure
CREAT_TX	Serum creatinine at time of transplant (Range: 27-400 μmol/l, Mean: 101.1 μmol/l, 1 missing)
INR_TX	INR at time of transplant (Range: 1-18, Mean: 1.5, 161 missing)
BILIRUBIN_TX	Serum bilirubin at time of transplant (Range: 2-1151 μmol/l, Mean: 96.4 μmol/l, 11 missing)
SODIUM_TX	Serum sodium at time of transplant (Range: 112-150 mmol/l, Mean: 136.2, 5 missing)
UKELD_TX	UKELD score at time of transplant (Range: 40.7-86.5, Mean: 55.0, 173 missing)
DSEX	Donor sex (Levels: 1 Male (n=1835), 2 Female (n=1663))
DWEIGHT	Donor weight at time of donation (Range: 22-140 kg, Mean: 74.3 kg, 6 missing)
DHEIGHT	Donor height at time of donation (Range: 52-208 cm, Mean: 170.7 cm, 34 missing)
DONOR_TYPE	Donor type (Levels: 1 Deceased Heartbeating (n=3321), 2 Deceased Non-Heartbeating (n=177))
DBG	Donor blood group (Levels: 0 O (n=1511), 1 A (n=1498), 2 B (n=384), 3 AB (n=105))
DCOD	Donor cause of death (See Table 1.8)
DCOD_GRP	Donor cause of death grouped (See Table 1.9)
DETHNIC	Donor ethnicity (Levels: 1 White (n=3332), 2 Asian or Asian-British (n=47), 3 Black or Black-British (n=35), 4 Chinese/Oriental (n=8), 6 Mixed (n=15), 7 Other (n=10), 9 Unknown (n=51))
DAGE	Donor age at time of donation (Range: 5-85 years, Mean: 44.8 years, 12 missing)

Table 1.2: The variables in the Liver Registration data set that are applicable only to patients who are transplanted. Details of these variables given, for continuous variables these are the range and mean and for factorial variables these are the levels of the factors and the number of observations with each level of the factor. Also the numbers of applicable patients who are missing values for these variables are given.

Code	Primary liver disease	Code	Primary liver disease
410	Chronic liver failure cause unknown	442	Hepatocellular carcinoma - cirrhotic
411	Primary biliary cirrhosis	443	Cholangiocarcinoma
412	Autoimmune chronic active liver disease	445	Secondary hepatic malignancy
413	Hepatitis B cirrhosis	447	Other primary hepatic malignancy, please specify
414	Primary sclerosing cholangitis	448	Benign liver tumour
415	Alpha-1-antitrypsin deficiency	450	Other metabolic liver disease, please specify
416	Budd-Chiari syndrome (not code 27)	451	Cystic fibrosis
417	Cryptogenic cirrhosis	460	Polycystic liver disease
418	Secondary biliary cirrhosis	461	Hereditary haemochromatosis
419	Alcoholic liver disease	462	Glycogen storage disease
420	Biliary atresia	471	Acute rejection
421	Congenital hepatic fibrosis	472	Chronic rejection
422	Wilson's Disease	473	Primary non-function
423	Congenital biliary disease	474	Acute vascular occlusion (artery plus vein)
424	Hepatitis C cirrhosis	475	Non-thrombotic infarction
425	Paediatric cholestatic liver disease, please specify	476	Ductopenic rejection
426	Non-alcoholic fatty liver disease	477	Recurrent disease
430	Acute hepatic failure - serologically indeterminate	478	Biliary complications
434	Acute hepatic failure - Wilson's disease	479	Hepatic artery thrombosis
436	Acute hepatic failure - HBV	480	Early graft dysfunction
437	Acute Hepatic Failure - paracetamol hepatotoxicity	482	Acute vascular occlusion - artery and venous
438	Acute Hepatic Failure - other drug toxicity	498	Other, please specify
439	Acute Hepatic Failure - other	499	Unknown
441	Hepatocellular carcinoma - non-cirrhotic	888	Not reported

Table 1.3: The codes used for different liver diseases

Level	Primary liver disease group
1	Primary biliary cirrhosis (PBC) (liv_dis=411, n=580)
2	Primary sclerosing cholangitis (PSC) (liv_dis=414, n=434)
3	Alcoholic liver disease (ALD) (liv_dis=419, n=1142)
4	Auto-immune + cryptogenic disease (AID) (liv_dis=412,417, n=523)
5	Hepatitis C cirrhosis (HCV) (liv_dis=424, n=686)
6	Hepatitis B cirrhosis (HBV) (liv_dis=413, n=162)
7	Cancers (liv_dis= 441,442,443,444,445,447, n=208)
8	Metabolic liver disease (liv_dis=415,422,426,434,450,456,461,462,466, 467,468, n=196)
9	Other liver diseases (liv_dis=410,416,418,420,421,423,425,427,448,451, 453,455,460,484,486,498,499,888, n=489)
10	Acute hepatic failure (liv_dis=428,430,435,436,437,438,439,471,472,473, 475,476,477,478,479,480,481,482,474,432, n=130)

Table 1.4: The groupings of the primary liver diseases given in Table 1.3

	Cause of death		Cause of death		Cause of death
0	Recipient still alive	532	Pulmonary infection (viral)	563	Bone marrow depression
500	Cause of death uncertain	533	Pulmonary infection (fungal)	564	Cachexia
511	Myocardial ischaemia and infarction	534	Infections elsewhere (except viral hepatitis)	566	Lymphoid malignant disease possibly induced by immunosuppressive therapy
512	Hyperkalaemia	535	Septicaemia	567	Lymphoid malignant disease not induced by immunosuppressive therapy
513	Haemorrhagic pericarditis	536	Tuberculosis (lung)	568	Malignant disease: lymphoproliferative disorders
514	Other causes of cardiac failure	537	Tuberculosis (elsewhere)	569	Dementia
515	Sudden unexplained cardiac death	538	Generalized viral infection	570	Sclerosing (or adhesive) peritoneal disease
516	Hypertensive cardiac failure	539	Peritonitis	571	Perforation of peptic ulcer
517	Hypokalaemia	541	Liver - due to hepatitis B virus	572	Perforation of colon
518	Fluid overload	542	Liver - other viral hepatitis	573	Non-lymphoid malignant disease possibly induced by immunosuppressive therapy
519	Elevated PVR/pulmonary hypertension	543	Liver - drug toxicity	574	Non-lymphoid malignant disease not induced by immunosuppressive therapy
520	Airway dehiscence	544	Cirrhosis - not viral	575	Early graft dysfunction
521	Pulmonary embolus	545	Cystic liver disease	576	Cardiac tamponade
522	Cerebro-vascular accident	546	Liver failure - cause unknown	577	ARDS
523	Gastro-intestinal haemorrhage	547	Renal failure	578	Respiratory failure
524	Haemorrhage from graft site	548	Recurrent primary disease - benign	579	Multi-system failure
525	Haemorrhage from vascular access or dialysis circuit	549	Recurrent primary disease - malignant	581	Accident related to treatment
526	Haemorrhage from ruptured vascular aneurysm	551	Patient refused further treatment	582	Accident unrelated to treatment
527	Haemorrhage from surgery	552	Suicide	590	Donor organ failure
528	Other haemorrhage	553	Therapy ceased for any other reason	595	Other identified cause of death
529	Mesenteric infarction	554	ESRF treatment withdrawn for medical reasons	598	Other identified cause of death
530	Pulmonary infection (protozoal)	561	Uraemia caused by graft failure	599	Unknown
531	Pulmonary infection (bacterial)	562	Pancreatitis	888	Cause of death not requested

Table 1.5: The codes for the primary cause of death for patients on the waiting list for a liver transplant

Code(s)	Cause of death group	Code(s)	Cause of death group
500	Cause of death uncertain	561	Uraemia caused by graft failure
511	Myocardial ischaemia and infarction	562	Pancreatitis
512	Hyperkalaemia	563	Bone marrow depression
513-518	Cardiac - miscellaneous	564	Cachexia
519	Elevated PVR	566-567	Lymphoma
520	Tracheal dehiscence	569	Dementia
521	Pulmonary embolus	570	Sclerosing (or adhesive) peritoneal disease
522	Cerebro-vascular accident	571	Perforation of peptic ulcer
523	Gastro-intestinal haemorrhage	572	Perforation of colon
524-528	Haemorrhage - miscellaneous	573-574	Non-lymphoid malignant disease
529	Mesenteric infarction	575	Early graft dysfunction
530-533	Pulmonary infection	576	Cardiac tamponade
534,536-539	Infection - miscellaneous	577	Ards
535	Septicaemia	578	Respiratory failure
541-546	Liver disease	579	Multi-system failure
547	Renal Failure (not kidney recipients)	581	Accident related to treatment
548	Recurrent primary disease - benign	582	Accident unrelated to treatment
549	Recurrent primary disease - malignant	590	Donor organ failure
551	Patient refused further treatment	595,598	Other identified cause of death
552	Suicide	599	Unknown
553	Therapy ceased for any other reason		

Table 1.6: The grouped causes of death for patients on the waiting list for a liver transplant

Code	Cause of graft failure	Code	Cause of graft failure
0	Graft still functioning	470	Recurrent disease
410	Acute rejection	480	Biliary complications
420	Chronic rejection	490	Recip. died, graft still functioning at T.O.D.
430	Primary non-functioning	495	Other
440	Acute vascular occlusion	498	Other, please specify
441	Vascular occlusion	499	Unknown
450	Non-thrombotic infarction		
460	Ductopenic rejection		

Table 1.7: The codes for causes of graft failure

Code	Cause of death	Code	Cause of death
0	Living donor	51	Pneumonia
10	Intracranial haemorrhage	52	Asthma
11	Intracranial thrombosis	53	Respiratory failure
12	Brain tumour	54	Carbon monoxide poisoning
13	Hypoxic brain damage - all causes	59	Respiratory - type unclassified (inc smoke inhalation)
19	Intracranial - type unclassified (CVA)	60	Cancer (other than brain tumour)
20	Trauma RTA - car	70	Meningitis
21	Trauma RTA - motorbike	71	Septicaemia
22	Trauma RTA - pushbike	72	Infections - type unclassified
23	Trauma RTA - pedestrian	73	Acute blood loss/hypovolaemia
29	Trauma RTA - unknown type	74	Liver failure (not self poisoning)
30	Other trauma - known or suspected suicide	75	Renal failure
		76	Multi-organ failure
31	Other trauma - accident	77	Sudden Infant Death Syndrome (SIDS)
39	Other trauma - unknown cause		
40	Cardiac arrest	80	Alcohol poisoning
41	Myocardial infarction	81	Paracetamol overdose
42	Aneurysm	82	Other drug overdose
43	Ischaemic heart disease	85	Self poisoning - type unclassified
44	Congestive heart failure	88	Not reported
45	Pulmonary embolism	90	Other identified cause of death
49	Cardiovascular - type unclassified	98	Other identified cause of death
		99	Unknown
50	Chronic pulmonary disease		

Table 1.8: The codes for donor cause of death

Code	Donor cause of death group
0	Live
10-11,19	CVA
12,13,40-45,49,50-54,59,60	Miscellaneous
70-77,80-82,85,88,90,98,99	Miscellaneous (continued)
20-23,29	RTA
30,31,39	Other trauma

Table 1.9: The donor cause of death groups

Initial data analysis

In this section, some initial data analysis on the Liver Registration data set is carried out assuming all the censoring in the data set is non-informative. Firstly, Kaplan-Meier estimates of the survival functions for time to death and time to censoring are obtained. Then the significant variables for both time to death and time to censoring are identified.

Figures 1.2 and 1.3 are plots of the Kaplan-Meier estimates of the survival functions for time to death and time to censoring respectively. The estimated median time to death is 1194 days and the estimated median time to censoring is 97 days. This shows that patients who are censored tend to spend less time on the list than those who die while on the waiting list.

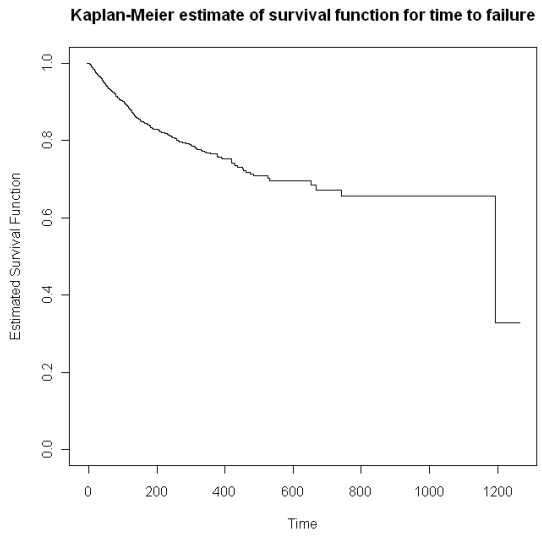

Figure 1.2: The Kaplan-Meier estimate of the survival function for time to failure

Tables 1.10 and 1.11 contain the results of Cox proportional hazards models for time to death and time to censoring for the Liver Registration data set. The general proportional hazards model is given by

$$h_i(t) = \exp(\beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_p x_{pi}) h_0(t),$$

where $\boldsymbol{\beta}$ is the vector of parameters of the explanatory variables $x_1, x_2, \ldots, x_p$ included in the model and $h_0(t)$ is the baseline hazard function. When fitting the proportional hazards model proposed by Cox (1972), no assumptions are made about the baseline hazard function and only $\boldsymbol{\beta}$ is estimated. The variables that are significant for time to death under the Cox proportional hazards model are

- UKELD score at time of registration,

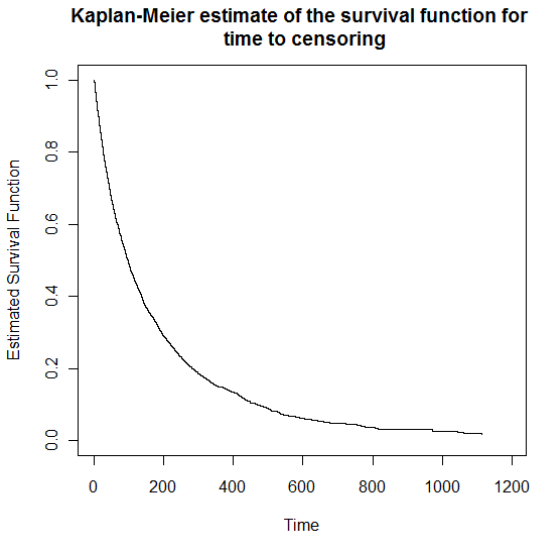

Figure 1.3: The Kaplan-Meier estimate of the survival function for time to censoring

- primary liver disease category,
- age at time of registration,
- ethnicity,
- serum sodium at time of registration and
- INR at time of registration.

The variables that are significant for time to censoring under the Cox proportional hazards model are

- UKELD score at time of registration,
- primary liver disease category,
- height at time of registration and
- blood group.

Parametric survival models will also be used in some of the methods reviewed in Chapter 3 so the exponential proportional hazards model is also fitted to the Liver Registration data set. This model has the form

$$h_i(t) = \exp(\mu + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_p x_{pi}),$$

where μ is the intercept parameter. The results of the exponential proportional hazards model for time to death and time to censoring are given in 1.12 and 1.13 respectively. The

variables that are significant for time to death under the exponential proportional hazards model are

- UKELD score at time of registration,
- primary liver disease category,
- age at time of registration and
- ethnicity.

The variables that are significant for time to censoring under the exponential proportional hazards model are

- UKELD score at time of registration,
- primary liver disease category,
- height at time of registration and
- blood group.

The results from the Cox proportional hazards models and exponential proportional hazards models are similar, as the models being fitted are both variations of the general proportional hazards model. The main difference is that the Cox proportional hazards model for time to death includes serum sodium and INR at time of registration as well. This suggests that some changes should be made to the UKELD score as these two variables are components of the UKELD score. But, as discussed in Section 1.2.1, we are using the original UKELD score which has since been revised, so this is not too surprising.

The four main covariates for time to death, UKELD score, age, primary liver disease category and ethnicity were examined, and it was found that only 2650 rows of data had full information for these covariates. If the two additional covariates that are significant for time to censoring, height and blood group, are also examined, then only 2605 rows of data had full covariate information. We will ignore any observations that do not have full covariate information. There are other methods of dealing with missing data that would be preferable but the aim here is to deal with the issue of informative censoring.

UKELD Score

As discussed in Section 1.2.1, the UKELD score used in our analyses is calculated using (1.3). However, this UKELD score has since been revised and is now given by (1.4). To asses how much change this causes in the UKELD score, the original value from (1.3) and the revised value from (1.4) are plotted against each other for each individual in the

Parameter	Estimate	Standard Error	p-value	Hazard Ratio
UKELD score	0.25126	0.01868	$< .0001$	1.286
PLD - PBC	-0.04365	0.34301	0.8987	0.957
PLD - PSC	-0.60813	0.38124	0.1107	0.544
PLD - ALD	-0.13828	0.32027	0.1864	0.871
PLD - AID	0.17176	0.33786	0.6112	1.187
PLD - HCV	0.59734	0.33487	0.0745	1.817
PLD - HBV	0.02856	0.50702	0.9551	1.029
PLD - Cancer	-1.09539	0.77467	0.1574	0.334
PLD - Metabolic	0.95969	0.36540	0.0086	2.611
PLD - Other	0.47124	0.34028	0.1661	1.602
PLD - Acute	0			
Age	0.02946	0.00587	$< .0001$	1.030
Ethnicity - White	1.26645	1.00791	0.2089	3.548
Ethnicity - Asian	0.34320	1.04198	0.7419	1.409
Ethnicity - Black	1.25277	1.09829	0.2540	3.500
Ethnicity - Chinese	-0.41759	1.43131	0.7705	0.659
Ethnicity - Other	0			
Serum Sodium	0.06060	0.01644	0.0002	1.062
INR	-0.22431	0.09654	0.0202	0.799

Table 1.10: Results for Cox model for time to death fitted to the Liver Registration data set assuming non-informative censoring

Parameter	Estimate	Standard Error	p-value	Hazard Ratio
UKELD score	0.02680	0.00471	$< .0001$	1.027
PLD - PBC	-0.05982	0.14472	0.6793	0.942
PLD - PSC	-0.03346	0.14420	0.8165	0.967
PLD - ALD	-0.24277	0.13282	0.0676	0.784
PLD - AID	-0.15159	0.14348	0.2907	0.859
PLD - HCV	0.08462	0.13807	0.5399	1.088
PLD - HBV	0.05300	0.17830	0.7663	1.054
PLD - Cancer	0.46530	0.15586	0.0028	1.592
PLD - Metabolic	-0.04300	0.16682	0.7966	0.958
PLD - Other	-0.32864	0.14746	0.0258	0.720
PLD - Acute	0			
Height	0.00953	0.00252	0.0002	1.010
Blood Group - O	-0.58504	0.11535	$< .0001$	0.557
Blood Group - A	-0.24982	0.11493	0.0297	0.779
Blood Group - B	-0.21245	0.12788	0.0967	0.809
Blood Group - AB	0			

Table 1.11: Results for Cox model for time to censoring fitted to the Liver Registration data set assuming non-informative censoring

Parameter	Estimate	Standard Error	p-value	Hazard Ratio
Intercept	-19.9758	1.2216	< 0.0001	2×10^{-9}
UKELD score	0.1854	0.0091	$< .0001$	1.204
PLD - PBC	-0.1346	0.3389	0.6913	0.874
PLD - PSC	-0.6234	0.3789	0.0999	0.536
PLD - ALD	-0.3441	0.3100	0.2671	0.709
PLD - AID	0.0113	0.3308	0.9727	1.011
PLD - HCV	0.4000	0.3263	0.2204	1.492
PLD - HBV	0.0276	0.5065	0.9566	1.028
PLD - Cancer	-1.4751	0.7661	0.0542	0.229
PLD - Metabolic	0.6556	0.3554	0.0650	1.926
PLD - Other	0.3188	0.3368	0.3439	1.375
PLD - Acute	0			
Age	0.0287	0.0059	< 0.0001	1.029
Ethnicity - White	1.1834	1.0082	0.2405	3.265
Ethnicity - Asian	0.2172	1.0445	0.8353	1.243
Ethnicity - Black	0.9603	1.1196	0.3910	2.612
Ethnicity - Chinese	-0.5219	1.4282	0.7148	0.593
Ethnicity - Other	0			

Table 1.12: Results for an exponential proportional hazards model for time to death fitted to the Liver Registration data set assuming non-informative censoring

Parameter	Estimate	Standard Error	p-value	Hazard Ratio
Intercept	-8.1525	0.5266	< 0.0001	0.0003
UKELD score	0.0312	0.0047	< 0.0001	1.032
PLD - PBC	-0.0115	0.1451	0.9368	0.989
PLD - PSC	0.0080	0.1447	0.9557	1.008
PLD - ALD	-0.2159	0.1334	0.1056	0.806
PLD - AID	-0.1179	0.1440	0.4129	0.889
PLD - HCV	0.1354	0.1387	0.3288	1.145
PLD - HBV	0.1004	0.1786	0.5741	1.106
PLD - Cancer	0.5487	0.1564	0.0004	1.731
PLD - Metabolic	-0.0114	0.1675	0.9460	0.989
PLD - Other	-0.3300	0.1485	0.0262	0.719
PLD - Acute	0			
Height	0.0108	0.0025	< 0.0001	1.011
Blood Group - O	-0.6175	0.1158	< 0.0001	0.539
Blood Group - A	-0.2588	0.1156	0.0251	0.772
Blood Group - B	-0.2077	0.1287	0.1066	0.812
Blood Group - AB	0			

Table 1.13: Results for an exponential proportional hazards model for time to censoring fitted to the Liver Registration data set assuming non-informative censoring

Liver Registration data set. This scatterplot can be seen in Figure 1.4. We can see that apart from a few individuals, the original UKELD score and the revised UKELD score are almost identical. Therefore our results should not be greatly affected by the use of the original UKELD score rather than the revised UKELD score.

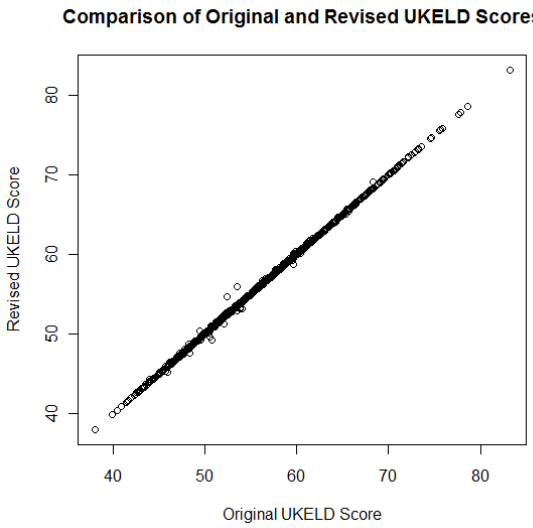

Figure 1.4: Scatterplot comparing the values of the original UKELD score given by (1.3) with the values of the revised UKELD score given by (1.4) for all individuals in the Liver Registration data set for whom the necessary component values are available.

1.2.6 Survival benefit

An important concept for the analysis of transplantation data is that of survival benefit. This quantifies the difference in survival between patients who received a transplant and similar patients who remained on the waiting list. From this it is possible to identify the patients who benefit most from liver transplantation and those that should remain on the waiting list at the present time. This approach was introduced in the USA by Merion et al. (2005), and subsequently there have been many analyses that use the concept of survival benefit.

The majority of the analyses use the covariate-adjusted hazard ratio for transplantation compared to not receiving a transplant to quantify the survival benefit of liver transplantation. Merion et al. (2005) use a time-dependent Cox regression model to calculate this hazard ratio. However, all the later analyses use the method of sequential stratification, introduced in Schaubel et al. (2006), which has been shown to give parameter estimates which can be more easily interpreted. As stated in Schaubel et al. (2008), sequential stratification is an "extension of Cox regression for evaluating time-dependent treatments

in the presence of time-dependent patient characteristics". This method creates a stratum each time a patient is transplanted and compares their survival with that of similar candidates who were active on the waiting list at the time. The experience from each stratum is aggregated to estimate the regression parameters using a Cox model.

The method is extended to deal with issues specific to liver transplantation, such as dependent censoring, in Schaubel et al. (2009b). This is covered in more detail in Chapter 6 where the method is applied to the Liver Registration data set. The contributions of each patient are weighted by the inverse of the probability of remaining untransplanted to account for the under-representation due to dependent censoring. This is one application of the well-established method known as inverse probability of censoring weighting (IPCW) which was introduced in Robins and Rotnitzky (1992) and is explained in more detail in Section 3.1.

It was established that overall, liver transplantation gives a significant survival benefit. However, patients have different severities of disease, as indicated by the MELD score, which is given in (1.2). The MELD score quantifies the level of organ dysfunction. A high MELD score indicates a high level of organ dysfunction. It has been shown by Merion et al. (2005) that survival benefit is not distributed evenly across subgroups of MELD scores. Those with high MELD scores have the greatest survival benefit from liver transplantation. In comparison, patients with fairly low MELD scores have a higher mortality risk post-transplant compared to remaining on the waiting list, and so they have a negative survival benefit from liver transplantation.

Donor factors should also be considered when computing the survival benefit as they effect the post-transplant mortality of recipients. The Donor Risk Index(DRI) can be used to measure the quality of the donor organ. Schaubel et al. (2008) carried out an analysis to compute the survival benefit for different levels of MELD scores and DRI. The patients with the highest MELD scores still receive a significant survival benefit, irrespective of the quality of the organ received. Those patients with a low MELD score who receive a high DRI liver have a significantly higher mortality risk than comparable patients who remain on the waiting list and possibly receive a better quality liver later. These results are especially worrying when the current organ allocation policy in the USA is considered. Patients with low MELD scores are generally given high DRI livers, so that the best quality organs can be given to the patients deemed to be the sickest. As shown by Volk et al. (2008), this has led to a small but significant decrease in the post-transplant survival of patients with low MELD scores.

More recently, analyses of the survival benefit for patients with specific diagnoses have been carried out in Lucey et al. (2009). Also the effect of individual components of the MELD score on survival benefit is computed in Sharma et al. (2009).

Using survival benefit to allocate donor livers One of the recent developments in the survival benefit literature is the application of survival benefit to the allocation of deceased-donor livers. This was first presented in Schaubel et al. (2009a). Here the definition of survival benefit is different from that given in the papers considered previously. This is to allow individual patients to be ranked in order of priority for a donor organ according to a benefit score. The benefit score is the candidate's 5-year mean lifetime with a transplant(from the organ under consideration) minus their 5-year mean lifetime without a transplant. An individual's 5-year mean lifetime is the area under the individual's survival curve out to 5 years, where the survival curve is found using a Cox model. In this approach, waiting list survival and post-transplant survival are modelled separately. Post-transplant survival is modelled using covariate-adjusted Cox regression. The model for waiting list survival is based on the sequential stratification approach. However, the paper giving the exact details of this method is yet to be published.

There have also been several articles that provide a critical analysis of the proposal to use survival benefit for donor liver allocation. The need for a high level of accuracy in the estimation of both pre- and post-transplant survival is highlighted by Kim and Kramers (2008). Some ethical questions are raised in Asrani et al. (2009); is it fair to give all patients with a certain diagnosis a lower priority just because a few will have a severe recurrence of the disease? Asrani at al. (2009) also discuss the limitations of the analyses that have been carried out so far. As all the analyses have been carried out on observational data, it is highly likely that there is a selection bias present in the data. This is because there are many factors that affect the matching of patients with donor organs that cannot be quantified in statistical modelling. Therefore even if an allocation policy that uses statistical models is implemented to assist with the selection of patients to be transplanted with a particular donor organ, the final decision on the suitability of a patient must belong to the clinicians.

1.3 Outline of Thesis

Chapters 2 and 3 provide a review of the methods in the literature that have been suggested to incorporate informative censoring into models. The most relevant methods are applied to the Liver Registration data set so that the results from the contrasting methods can be compared. Chapter 2 looks at estimators of the survival function that are used to give bounds on the possible values of the estimated survival function assuming informative censoring. The estimators in this chapter are some of the first estimators presented in the literature to allow for informative censoring. As the estimators in Chapter 2 give bounds

that are too wide to be of practical use, we review some of the more recent methods from the literature in Chapter 3. These can be split into two categories: estimators that use models of the censoring process and sensitivity analyses. The most popular of the methods in the former category is reweighting the estimators, particularly inverse probability of censoring weighted estimators. There are sensitivity analyses presented in Chapter 3 for both the Cox model and parametric survival models. These methods estimate the change in the parameter estimates in the model if informative censoring is assumed instead of non-informative censoring. The methods that use parametric models are less computationally intensive but lack the flexibility of the methods that use the Cox model.

Chapters 4 and 5 show the development of a new sensitivity analysis methodology that can generally be applied to any situation where there is potentially informative censoring. It uses piecewise exponential models, which means the method is computationally simple but more flexible than the method that uses standard parametric models. It is an extension of one of the sensitivity analysis methods detailed in Chapter 3. The derivation of the method is shown in Chapter 4, along with its application to the Liver Registration data set. A simulation study is carried out to test the accuracy of this new methodology and this is detailed in Chapter 5. This allows us to identify the situations where the sensitivity analysis is least accurate. An extension of the sensitivity analysis is presented to try and overcome these identified limitations.

Finally, in Chapter 6, a method that is of interest to NHSBT is detailed, that allows the survival benefit of groups of patients of interest to be calculated by comparing survival on the waiting list with survival after transplantation. This method has already been applied to US data, but we suggest some changes to the method and then apply it to the Liver Registration data set. The method overcomes any potentially informative censoring in the data set by using inverse probability of censoring weighted estimators, which are described in Chapter 3. There will also be a concluding chapter that summarises the main findings and shortcomings of this work and also describes further work that could be carried out.

Chapter 2

Bounds on the Marginal Survival Function under Informative Censoring

The first approaches that account for potentially informative censoring in data sets derive estimators that are extensions of the Kaplan-Meier estimator. Because of this they cannot incorporate covariates unless the variables have a very simple structure. Generally these estimators are used to provide bounds on the marginal survival function. These bounds tend to be quite wide despite efforts to derive tighter bounds.

Here a review of these methods is presented, in the order that they were published, so that it is possible to see the improvements in the methods. The estimators that are considered suitable are applied to the Liver Registration data set so that results obtained using the different methods can be contrasted. As this is a review chapter, all the methods discussed can be found in the literature and are presented here in consistent notation. Unless otherwise stated, the original work in this chapter is the application of the methods to the Liver Registration data set.

The estimators here use a variety of assumptions about the conditional distribution of the failure time variable given the censoring time variable to make the joint distribution of the two variables identifiable. These range from non-parametric methods to using a copula to specify the joint distribution of the variables. A section on copulas is included, detailing some of the more common forms used.

2.1 Measuring Dependence between Variables

When estimators are being compared, it is important that they assume the same amount of dependence between the time to censoring and time to failure variables so that meaningful

comparisons can be made. As the estimators use different parameters to control the dependence then, where possible, the relationship between these parameters and a widely used measure of dependence between two variables should be established. The measure of dependence that has been chosen for this is Kendall's τ which is a measure of concordance. The definition presented here is taken from Nelsen (1999). A pair of random variables are said to be concordant if large values of one variable are associated with large values of the other variable, and small values of one are associated with small values of the other. So, if there are two observations (x_i, y_i) and (x_j, y_j) from the random vector (X, Y), then they are concordant if $x_i < x_j$ and $y_i < y_j$, or if $x_i > x_j$ and $y_i > y_j$. Similarly, they are discordant if $x_i < x_j$ and $y_i > y_j$, or if $x_i > x_j$ and $y_i < y_j$. If $x_i = x_j$ or $y_i = y_j$, then the pair is neither concordant nor discordant.

To be able to express the concordance measure, firstly a concordance function, Q, needs to be defined. This is the difference of the probability of concordance and the probability of discordance between two vectors (X_1, Y_1) and (X_2, Y_2) with joint distribution functions H_1 and H_2, but common margins. So,

$$Q(H_1, H_2) = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0]. \tag{2.1}$$

The population version of Kendall's τ for random vectors (X_1, Y_1) and (X_2, Y_2), is the concordance function Q, given in (2.1), but assuming the same joint distribution function H for each of the random vectors. So, τ can be expressed as

$$\tau = Q(H, H) = 4 \iint_{\Re^2} H(x, y) dH(x, y) - 1,$$

as shown in Nelson (1999).

So that sensible values of Kendall's τ to be used here can be established, the relationship between this and the parameter δ, which we introduce in Section 3.3, will be found. The parameter δ is used here because the sensitivity analysis that uses this parameter is considered in detail in Chapter 4, so a sensible range of values for δ is established. We find that $\tau = \delta/2$ when using an approximation to the joint density function. When fitting the model that incorporates informative censoring to the Liver Registration data set, the 95% confidence interval obtained for δ is (0.1388,0.4163), so $\tau = 0.2$ will be used as the upper limit for Kendall's τ here. However, the dependence assumption that was used to obtain this interval for δ cannot be checked so it is possible that a larger value of Kendall's τ should be used.

For some of the estimators presented here it is not possible to relate their parameter directly to Kendall's τ. In these cases a parameter value that gives an estimator with the same median value as an estimator for which the parameter can be directly related to Kendall's τ is chosen.

2.2 Fisher-Kanarek Estimator

Fisher and Kanarek (1974) present a non-parametric method that estimates the survival function without consideration of covariates. The method allows for both informative and non-informative censoring within the same data set. In the case of no informative censoring the Kaplan-Meier estimate of the survival curve results.

The following model presented is non-parametric with the exception of the parameter α which either expands or contracts the residual lifetime after informative censoring. The assumption of the dependence between survival time T and informative censoring time C_I used is

$$P(T > t|C_I = c < t) = P(T > c + \alpha(t - c)|C_I > c + \alpha(t - c)). \tag{2.2}$$

If $\alpha > 1$, then it is the patients with a poorer prognosis who are censored, and $\alpha < 1$ means it is the patients with a favourable prognosis who drop out. If $\alpha = 1$ then censoring has no effect on expected survival and corresponds to the independent censoring case. This means the Kaplan-Meier estimate will be appropriate.

As there are three possible times that can be observed for each of the $i = 1, 2, \ldots, n$ individuals, there are three variables:

- $\tilde{T}$, the survival time if it is less than the censoring time, which has survival function $P(\tilde{T} > t) = S_{\tilde{T}}(t)$
- C_I, the censoring time that shall be considered informative, where individuals are lost to follow-up, which has survival function $P(C_I > t) = S_{C_I}(t)$, and
- C_E, the censoring time that will be considered independent of failure time, such as end of study censoring, which has survival function $P(C_E > t) = S_{C_E}(t)$.

The survival function $S(t)$ of the "true" survival time T is the function that will be estimated here. The survival time, T, will be equal to $\tilde{T}$ if $\tilde{T} \leq C_I$, otherwise the assumption in (2.2) is used. This means that $S(t)$ is related to $S_{\tilde{T}}(t)$ and $S_{C_I}(t)$ by the following relationship

$$\begin{aligned} S(t) &= P(T > t|C_I > t)P(C_I > t) \\ &\quad + \int_0^t P(T > t'|C_I = t')P(T > t|T > t', C_I = t')dP(C_I \leq t') \\ &= S_{\tilde{T}}(t)S_{C_I}(t) + \int_0^t S_{\tilde{T}}(t' + \alpha(t - t'))d(-S_{C_I}(t')). \end{aligned} \tag{2.3}$$

So to estimate the survival function $S(t)$, estimates of $S_{\tilde{T}}(t)$ and $S_{C_I}(t)$ need to be obtained and then substituted into (2.3).

The data observed are $Y_i = \min(\tilde{T}_i, C_{Ii}, C_{Ei})$ and the indicator functions

$$\Delta_{i,\tilde{T}} = \begin{cases} 1, & \text{if } \tilde{T}_i < \min(C_{Ii}, C_{Ei}) \\ 0, & \text{otherwise} \end{cases}$$

$$\Delta_{i,C_I} = \begin{cases} 1, & \text{if } C_{Ii} \leq \tilde{T}_i, C_{Ii} < C_{Ei} \\ 0, & \text{otherwise} \end{cases}$$

and

$$\Delta_{i,C_E} = \begin{cases} 1, & \text{if } C_{Ei} < \min(\tilde{T}_i, C_{Ii}) \\ 0, & \text{otherwise.} \end{cases}$$

Let $Y_{(i)}$ be the corresponding order statistics and $\Delta_{(i),\tilde{T}}$, $\Delta_{(i),C_I}$ and $\Delta_{(i),C_E}$ the indicator functions relating to these order statistics. The maximum likelihood estimates of $S_{\tilde{T}}(t)$, $S_{C_I}(t)$ and $S_{C_E}(t)$ are then given by

$$\begin{aligned} \hat{S}_{\tilde{T}}(t) &= \prod_{i=1,2,\ldots,k} \left(\frac{n-i}{n-i+1}\right)^{\Delta_{(i),\tilde{T}}} \quad \text{where} \quad Y_{(k)} \leq t < Y_{(k+1)} \\ \hat{S}_{C_I}(t) &= \prod_{i=1,2,\ldots,k} \left(\frac{n-i}{n-i+1}\right)^{\Delta_{(i),C_I}} \quad \text{where} \quad Y_{(k)} \leq t < Y_{(k+1)} \\ \text{and} \quad \hat{S}_{C_E}(t) &= \prod_{i=1,2,\ldots,k} \left(\frac{n-i}{n-i+1}\right)^{\Delta_{(i),C_E}} \quad \text{where} \quad Y_{(k)} \leq t < Y_{(k+1)}. \end{aligned} \tag{2.4}$$

So the maximum likelihood estimate of $S(t)$ is given by

$$\hat{S}(t) = \hat{S}_{\tilde{T}}(t)\hat{S}_{C_I}(t) + \int_0^t \hat{S}_{\tilde{T}}(t' + \alpha(t-t'))d(-\hat{S}_{C_I}(t')) \tag{2.5}$$

where $\hat{S}_{\tilde{T}}(t)$ and $\hat{S}_{C_I}(t)$ are the product-limit estimates defined in (2.4).

As the data give no information about the value of α, assumed values of α should be used. These can be used to see how robust the assumption of non-informative censoring is. If a large value of α is used, then the true marginal survival distribution should lie somewhere in the region between the Fisher-Kanarek estimator and the Kaplan-Meier estimator. However, this is not guaranteed as this method does not provide true bounds on the marginal survival function.

2.3 Peterson Bounds on the Survival Function

Methods that give definite bounds on the survival function are now considered, but these methods only allow for one type of censoring. So even if we have end of study censoring, it has to be treated as possibly informative censoring in these methods.

Peterson (1976) gives bounds for a joint survival function $G(t_1, t_2) = P(T > t_1, C > t_2)$ and the marginal survival functions $S_T(t)$ and $S_C(t)$. The estimated survival function of the variable T is of interest, so the bounds for the marginal survival function $S_T(t)$ are obtained. These are given by

$$S_T^*(t) + S_C^*(t) \leq S_T(t) \leq S_T^*(t) + (1 - p_1) \tag{2.6}$$

where

$$\begin{aligned} S_T^*(t) &= P(T > t, T < C), \\ S_C^*(t) &= P(C > t, C < T) \\ \text{and} \quad 1 - p_1 &= 1 - P(T < C) = P(C < T). \end{aligned}$$

The observed data are $Y_i = \min(T_i, C_i)$ and $\Delta_i = I(Y_i = T_i)$ for $i = 1, 2, \ldots, n$. The empirical estimators of the marginal survival functions are used as they are consistent estimators, where

$$\begin{aligned} \hat{S}_T^*(t) &= \frac{1}{n} \sum_{i=1}^{n} I[Y_i \geq t, \Delta_i = 1] \\ \text{and} \quad \hat{S}_C^*(t) &= \frac{1}{n} \sum_{i=1}^{n} I[Y_i \geq t, \Delta_i = 0]. \end{aligned}$$

The empirical estimator of $(1 - p_1)$ is also used, which is given by

$$1 - \hat{p}_1 = \frac{1}{n} \sum_{i=1}^{n} I[\Delta_i = 0].$$

If these terms are substituted into (2.6) then consistent estimators of the bounds for $S_T(t)$ can be obtained. After a little algebra they become

$$\frac{1}{n} \sum_{i=1}^{n} I[Y_i \geq t] \leq S_T(t) \leq \frac{1}{n} \sum_{i=1}^{n} I[Y_i \geq t] + \frac{1}{n} \sum_{i=1}^{n} I[Y_i < t, \Delta_i = 0]. \tag{2.7}$$

2.4 Slud-Rubinstein Bounds on the Survival Function

Slud and Rubinstein (1983) also derive bounds for the survival function $S_T(t)$ but their bounds can be tighter than those given by Peterson (1976). They make a nonparametric assumption on the joint density $f(t, c)$ of (T, C),

$$\lim_{\delta \to 0} \frac{P(t < T < t + \delta \,|\, T > t, C \leq t)}{P(t < T < t + \delta \,|\, T > t, C > t)} = \rho(t) \tag{2.8}$$

where $\rho(.)$ is a known function of t. This means that $\rho(t)$ is the amount that the conditional death hazard at time t differs by, according to whether the individual is censored before

or after t. So $\rho = 1$, corresponds to the independence assumption. If $\rho(t) > 1$ for all t, then there is positive dependence between T and C and similarly if ρ is always below 1, then there is negative dependence between failure and censoring.

If we assume that $\rho(.)$ is known, then there is a consistent estimator of the marginal survival function $S_T(t)$ which generalises the Kaplan-Meier estimator. Again $Y_i = \min(T_i, C_i)$ and $\Delta_i = I(Y_i = T_i)$ for $i = 1, 2, \ldots, n$ are observed. Let $Y_{(1)} \leq \ldots \leq Y_{(d)}$ be the ordered failure times, when there are d observations with $\Delta_i = 1$. Let the number of observations censored between $Y_{(j)}$ and $Y_{(j+1)}$ be c_j, with c_0 censored before the first failure time. The number of individuals with $Y_i \geq Y_{(j)}$ is defined to be n_j.

The product-limit estimator for $S_T(t)$ proposed in Slud and Rubinstein (1983) is

$$\hat{S}_\rho(t) = \frac{1}{n}\left\{ n(t) + \sum_{k=0}^{d(t)-1} c_k \prod_{i=k+1}^{d(t)} \frac{n_i - 1}{n_i + \rho_i - 1} \right\}, \tag{2.9}$$

where

$$n(t) = \sum_i I(Y_i \geq t), \quad d(t) = \sum_i I(Y_i \leq t, \Delta_i = 1) \quad \text{and} \quad \rho_i = \rho(Y_{(i)}).$$

After some algebra, this becomes

$$\hat{S}_\rho(t) = \prod_{i=1}^{d(t)} \frac{n_i - 1}{n_i + \rho_i - 1} + \frac{1}{n} \sum_{k=1}^{d(t)} (\rho_k - 1) \prod_{i=k}^{d(t)} \frac{n_i - 1}{n_i + \rho_i - 1}. \tag{2.10}$$

When $\rho(.) = 1$, $\hat{S}_\rho$ is exactly the Kaplan-Meier estimator.

From (2.9), we see that for fixed t, $\hat{S}_\rho(t)$ is a decreasing function of ρ, so as ρ increases, the value of $\hat{S}_\rho$ at time t decreases. Bounds for the function $\rho(.)$, as defined as in (2.8), can be assumed and if the true value of the function does lie between the bounds $\rho_1(.)$ and $\rho_2(.)$, then for sufficiently large samples

$$\hat{S}_{\rho_2}(t) \leq S(t) \leq \hat{S}_{\rho_1}(t). \tag{2.11}$$

This can be used to give bounds on the survival function which are tighter than those given in Peterson (1976), which correspond to (2.11) with $\rho_1 = 0$ and $\rho_2 = \infty$. However, as there is no information available on the value of ρ from the observed data, it is not possible to identify whether the bounds assumed contain the true value of ρ. This limits the usefulness of this method in a practical setting.

2.5 Klein-Moeschberger Bounds on the Survival Function

Klein and Moeschberger (1988) also present bounds on the survival function $S_T(t)$ that are tighter than those of Peterson (1976). However they make a different assumption about the

dependence between the failure and censoring times. As previously, the marginal survival functions of T and C are $S_T(t)$ and $S_C(t)$ respectively. The joint survival function of T and C, $G(t_1, t_2) = P(T > t_1, C > t_2)$ is expressed as

$$G(t_1, t_2) = \left\{ \left[\frac{1}{S_T(t_1)} \right]^{\theta-1} + \left[\frac{1}{S_C(t_2)} \right]^{\theta-1} - 1 \right\}^{\frac{-1}{\theta-1}} \tag{2.12}$$

for $\theta \geq 1$. This model for the joint distribution of T and C was first introduced in Clayton (1978) to model association in bivariate lifetimes. It is also used to model bivariate survival data in Oakes (1982). The model in (2.12) can be interpreted in terms of the hazard functions

$$h_T(t|C = c) = \lim_{\delta \to 0} \left[\frac{P(t \leq T < t + \delta | C = c, T \geq t)}{\delta} \right]$$

and

$$h_T(t|C > c) = \lim_{\delta \to 0} \left[\frac{P(t \leq T < t + \delta | C > c, T \geq t)}{\delta} \right].$$

Using (2.12), we obtain

$$h_T(t|C = c) = \theta h_T(t|C > c). \tag{2.13}$$

This means that for $\theta > 1$, the hazard rate of death, if censoring happens at time c, is the hazard rate of death if censoring had not occurred multiplied by θ. So the hazard rate if censoring does occur will be greater than the hazard rate if censoring does not occur, as it has been accelerated by a factor of θ. Therefore this only allows for positive dependence between T and C. This is not a problem for the data set under consideration here, as it is suspected that the dependence between T and C is positive.

Klein and Moeschberger (1988) show that $\tau = (\theta - 1)/(\theta + 1)$. Since $\theta \geq 1$, τ can only take values between 0 and 1.

If T and C have joint survival function (2.12), then the observed value $Y = \min(T, C)$ has survival function

$$F(t) = \left\{ \left[\frac{1}{S_T(t)} \right]^{\theta-1} + \left[\frac{1}{S_C(t)} \right]^{\theta-1} - 1 \right\}^{\frac{-1}{\theta-1}}.$$

This is a reasonable choice for the joint distribution function of T and C as it is used to model bivariate survival data in Oakes (1982). It is also related to the Clayton copula function given in Table 2.1, which seems to be a reasonable choice of copula family, as we will discuss in Section 2.10.1. The marginal distribution function of T is also required, which is defined as

$$Q_1(t) = P(Y < t, T < C).$$

These functions are estimated directly from the observed data, $Y_i = \min(T_i, C_i)$ and $\Delta_i = I(Y_i = T_i)$ for $i = 1, 2, \ldots, n$, using

$$\hat{F}(t) = \sum_{i=1}^{n} \frac{I(Y_i \geq t)}{n} \quad \text{and} \quad \hat{Q}_1(t) = \sum_{i=1}^{n} \frac{I(Y_i \leq t, \Delta_i = 1)}{n}.$$

Then, if θ is known and the underlying joint survival of (T, C) is given by (2.12), a consistent estimator of $S_T(t)$, given in Klein and Moeschberger (1988), is $\hat{S}_\theta(t)$ where

$$\hat{S}_\theta(t) = \begin{cases} \left\{ 1 + (\theta - 1) \int_0^t \frac{d\hat{Q}_1(u)}{[\hat{F}(u)]^\theta} \right\}^{\frac{-1}{\theta - 1}} & \text{if } \theta > 1 \\ \exp\left[- \int_0^t \frac{d\hat{Q}_1(u)}{\hat{F}(u)} \right] & \text{if } \theta = 1. \end{cases} \tag{2.14}$$

Upper and lower bounds on $S_T(t)$ can be found by letting $\theta \to 1^+$ and $\theta \to \infty$ respectively. This gives an upper bound which corresponds to independence between T and C and a lower bound which is the same as that of both Peterson (1976) and Slud and Rubinstein (1983). So,

$$F(t) \leq S_T(t) \leq \exp\left[- \int_0^t F^{-1}(u) dQ_1(u) \right].$$

It is possible to set up tighter bounds on the survival function using the same method as Slud and Rubinstein (1983). A possible range of values for θ, (θ_1, θ_2) is specified. If the sample size is sufficiently large and $\theta_1 \leq \theta \leq \theta_2$, then $\hat{S}_{\theta_1}(t) \geq S(t) \geq \hat{S}_{\theta_2}(t)$.

2.6 Applying Methods to Liver Registration Dataset

Firstly, the Fisher-Kanarek estimator is fitted to the Liver Registration data set. As it is believed that the transplant candidates with the poorest prognosis that are being censored, then it is assumed that α is greater than 1 when obtaining an estimate of the survival function. More specifically, $\alpha = 3$ is chosen. This is the value of α that gives the estimator the same median value as the Klein-Moeschberger estimator for this data set with Kendall's $\tau = 0.2$. The estimate for the independent case ($\alpha = 1$), which is the Kaplan-Meier estimate of the survival function, is also obtained to see how close the other estimates are. These are given in Figure 2.1. From this we see that as α increases, the estimate of the survival function decreases more quickly. This is expected as we are adjusting for patients who survive for progressively shorter times after censoring. Also, there can be a rather large difference between the Fisher-Kanarek estimate of the survival function and the Kaplan-Meier estimate.

Drawbacks of this method When the largest observation time in a data set, t^*, is censored, the Kaplan-Meier estimate of the survival function cannot be defined beyond this time. As stated in Kaplan and Meier (1958), the estimated survival function beyond

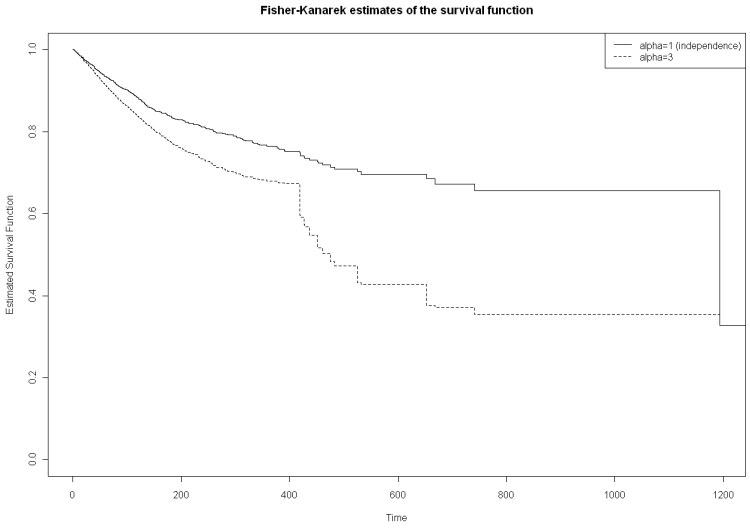

Figure 2.1: Plot of Fisher-Kanarek estimators of the marginal survival function for T for $\alpha = 1$ and 3

time t^* will lie somewhere between $\hat{S}(t^*)$ and 0, but it is not possible to define it any more precisely. However, Fisher and Kanarek have disregarded this when presenting their estimates of the survival function. In their simulated data set, they observed 44 deaths with times between .01 and .58, 34 non-informative censoring times between .01 and .67 and 22 informative censoring times between .03 and .72. Therefore, their last observation must be censored. So, the product-limit estimate of $S_{\tilde{T}}$ should not be defined beyond .72. Fisher and Kanarek do not explicitly state that they assume a value for $S_{\tilde{T}}$ beyond this time, but they present estimates of $S(t)$ up until time 1 when $\alpha > 1$. This means that they must have defined $S_{\tilde{T}}$ beyond .72 so that the integral in (2.5) can be evaluated.

In the Liver Registration data set, the last observation is a censored one at time 1265. This means that here $S_{\tilde{T}}$ should not be defined beyond this time. However, we assume that beyond time 1265, $S_{\tilde{T}}$ remains at the same value that it has at time 1265. This is why we observe the strange behaviour of the Fisher-Kanarek estimator at around time 400 in Figure 2.1. It is for this reason that use of this estimator is not recommended.

The Slud-Rubinstein bounds for the survival function estimate for the Liver Registration data set are shown in Figure 2.2. A range of values for ρ were chosen, with the upper and lower bounds being $\rho = 0$ and $\rho = \infty$ respectively. As these are the same as the Peterson bounds on the survival function, it was not necessary to produce a separate plot for these bounds. We chose $\rho = 2.7$ as this gives an estimator with the same median as the Klein-Moeschberger estimator with $\tau = 0.2$. The estimator with $\rho = 1$ is included as this is the same as the Kaplan-Meier estimator and can be used for comparison. We see

that the Peterson bounds are extremely wide and the Slud-Rubinstein bounds can indeed be an improvement on these. However, depending on our confidence about the bounds on the value of ρ, the Slud-Rubinstein bounds may still be wide and so are of little practical value.

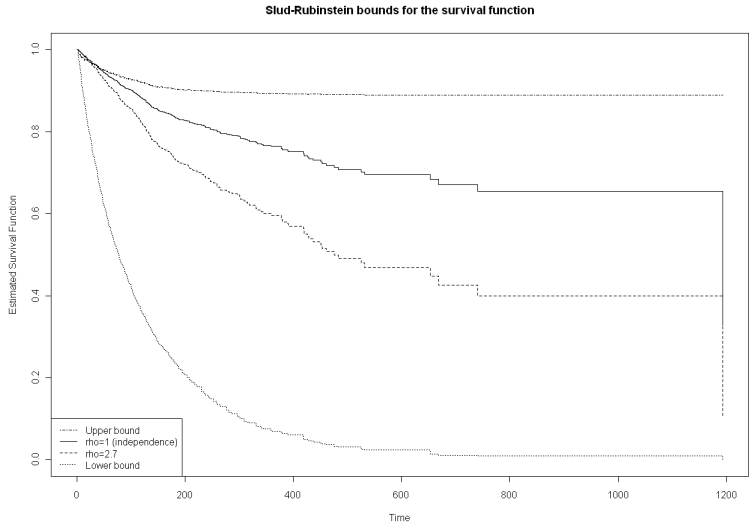

Figure 2.2: Plot of Slud-Rubinstein bounds on the marginal survival function for T and a Slud-Rubinstein estimator of the marginal survival function for T with $\rho = 2.7$.

Figure 2.3 shows the Klein-Moeschberger bounds for the survival function for the Liver Registration data set. Estimators for values of θ between 1 and ∞ are presented. So, the lower bound is the same as that for Peterson and Slud-Rubinstein but the upper bound corresponds to the assumption of independence. Therefore, the Klein-Moeschberger bounds are not as wide as those of Peterson. It also means that only positive correlation between the failure and censoring times is considered. While this is not a problem for this particular data set, the method may not be suitable for other data sets. The estimator with $\theta = 1.5$ corresponds to $\tau = 0.2$.

2.7 Background on Copulas

Copulas can be used to give the dependence structure between two variables X and Y with marginal distribution functions F and G respectively. All the definitions and information on copulas given in this section come from Nelsen (1999). It is possible to define either the joint distribution function H, or the joint survival function $\bar{H}$. Firstly, the definition that uses the joint distribution function which comes from Sklar's theorem is given.

Sklar's Theorem *If we have a joint distribution function H with margins F and G, then*

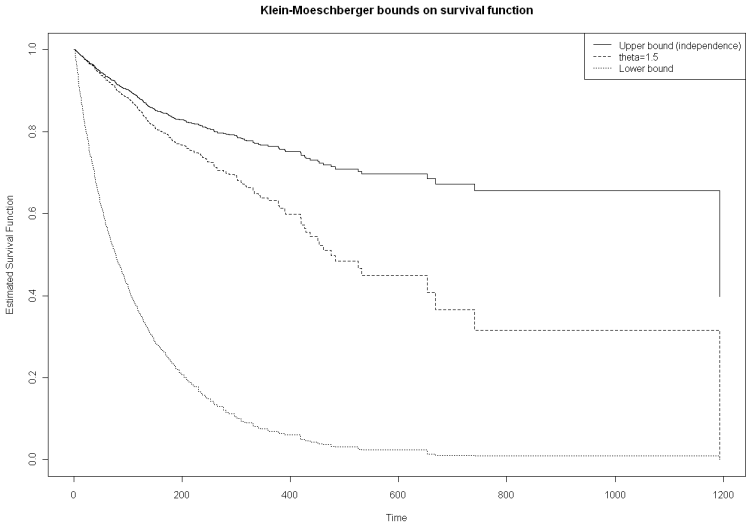

Figure 2.3: Plot of Klein-Moeschberger bounds on the marginal survival function for T and a Klein-Moeschberger estimator of the marginal survival function for T with $\theta = 1.5$.

there exists a copula $\mathcal{C}$ *such that for all* x, y *in* $\mathbb{R}$

$$H(x, y) = \mathcal{C}(F(x), G(y)). \tag{2.15}$$

This means that the copula is given by

$$\mathcal{C}(u, v) = H(F^{-1}(u), G^{-1}(v)), \quad u, v \in [0, 1],$$

where F^{-1} and G^{-1} are inverses of F and G.

2.7.1 Survival copulas

A similar function $\bar{\mathcal{C}}$, called a survival copula, can be defined. This gives the joint survival function $\bar{H}$ in terms of the marginal survival functions $\bar{F}$ and $\bar{G}$. Again from Nelsen (1999), we have

$$\bar{H}(x, y) = \bar{\mathcal{C}}(\bar{F}(x), \bar{G}(y)).$$

This is related to the copula defined in (2.15) by

$$\bar{\mathcal{C}}(u, v) = u + v - 1 + \mathcal{C}(1 - u, 1 - v).$$

This relationship is obtained since

$$\begin{aligned}
\bar{H}(x, y) &= 1 - F(x) - G(y) + H(x, y) \\
&= \bar{F}(x) + \bar{G}(y) - 1 + \mathcal{C}(F(x), G(y)) \\
&= \bar{F}(x) + \bar{G}(y) - 1 + \mathcal{C}(1 - \bar{F}(x), 1 - \bar{G}(y)).
\end{aligned}$$

2.7.2 Archimedean copulas

There is a special class of copula functions known as Archimedean copulas, defined in Nelsen (1999), where the copula can be expressed as

$$\mathcal{C}(u,v) = \varphi^{-1}(\varphi(u) + \varphi(v)), \tag{2.16}$$

where φ is the generator of the copula. The most well-known one parameter families of Archimedean copulas are given in Table 2.1.

Name	$\mathcal{C}(u,v)$	$\varphi(t)$	$\theta \in$
Clayton	$\max([u^{-\theta} + v^{-\theta} - 1]^{-1/\theta}, 0)$	$\frac{1}{\theta}(t^{-\theta} - 1)$	$[-1,\infty)\backslash\{0\}$
Gumbel-Hougaard	$\exp(-[(-\log u)^{\theta} + (-\log v)^{\theta}]^{1/\theta})$	$(-\log t)^{\theta}$	$[1,\infty)$
Frank	$-\frac{1}{\theta}\log(1 + \frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1})$	$-\log(\frac{e^{-\theta t}-1}{e^{-\theta}-1})$	$(-\infty,\infty)\backslash\{0\}$

Table 2.1: Table showing some of the most well-known families of Archimedean copulas with their generators and corresponding copula functions as given in Nelsen (1999).

Some of the papers considered in this section refer to a gamma frailty copula, which is given by

$$\mathcal{C}(u,v) = u + v - 1 + \left[(\frac{1}{1-u})^{\alpha-1} + (\frac{1}{1-v})^{\alpha-1} - 1\right]^{-\alpha/(\alpha-1)}, \quad \alpha > 0\backslash\{1\}.$$

It is easy to see that this is the corresponding survival copula for the Clayton copula with $\theta = \alpha - 1$. However, in some papers the domain of α is restricted to $(1,\infty]$ so that only positive dependence between the variables is possible.

To visualise the differences between the copula functions considered, the density function of the copula, $c(u,v) = \frac{\partial^2}{\partial u \partial v}\mathcal{C}(u,v)$, for the Clayton, Frank, Gumbel-Hougaard and gamma frailty copulas has been plotted in Figure 2.4.

2.7.3 A dependence measure for copulas

Kendall's τ has already been defined as a measure of concordance and it will be used to express the amount of dependence between time to failure and time to censoring here. It will be shown that it can be expressed in terms of copulas, instead of the joint distribution function. Recall that the population version of Kendall's τ for random vectors (X_1, Y_1) and (X_2, Y_2), each with joint distribution function H is

$$\tau = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0].$$

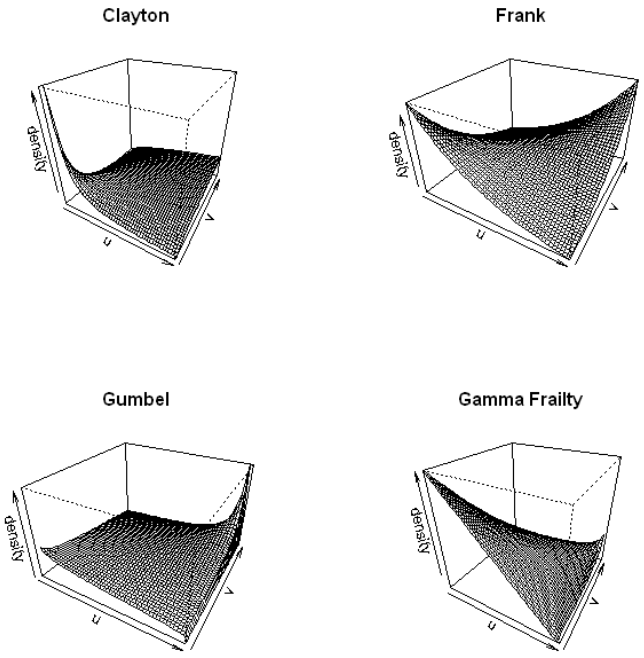

Figure 2.4: Joint density functions for Clayton, Frank, Gumbel-Hougaard and gamma frailty copulas

Let X and Y be continuous random variables with copula $\mathcal{C}$, then the population version of Kendall's τ given in Nelsen (1999) is

$$\tau_{\mathcal{C}} = Q(\mathcal{C}, \mathcal{C}) = 4 \iint_{\mathbb{I}^2} \mathcal{C}(u, v) d\mathcal{C}(u, v) - 1.$$

where the unit square $\mathbb{I}^2$ is the product $\mathbb{I} \times \mathbb{I}$ where $\mathbb{I} = [0, 1]$. For an Archimedean copula with generator φ, this has a simpler form, given by

$$\tau_{\mathcal{C}} = 1 + 4 \int_0^1 \frac{\varphi(t)}{\varphi'(t)} dt.$$

The following theorem from Georges et. al. (2001) will be of use when using copulas that are the corresponding survival copulas of well-known families of copulas.

Theorem *The Kendall's τ of the survival copula $\hat{\mathcal{C}}$ are equal to the Kendall's τ of the associated copula $\mathcal{C}$*

2.8 Self-consistent Estimators based on an Assumed Copula

In this section, the estimator from Zheng and Klein (1994) is presented, which they call a self-consistent estimator based on an assumed copula. The idea of self-consistency was first discussed in Efron (1976), and a summary of this concept is given here.

If both the time to the event of interest T, and the time to censoring C could be observed for every individual in a data set, then natural non-parametric estimators of the marginal survivor functions, $S_T(t)$ and $S_C(t)$, of T and C respectively, would be

$$\hat{S}_T(t) = \frac{1}{n} \sum_{i=1}^{n} I[T_i \geq t] \quad \text{and} \quad \hat{S}_C(t) = \frac{1}{n} \sum_{i=1}^{n} I[C_i \geq t].$$

As we have censored data, these estimators need to be adapted. Let $Y_1, Y_2, \ldots, Y_n$ be the observation times, where $Y_i = \min(T_i, C_i)$. If Y_i is a death time then it is known whether T_i is smaller or greater than t. If Y_i is a censored observation that is greater than or equal to t, then it is also known that the T_i for this individual is greater than t. However, if Y_i is a censored observation that is less than t, it is not known if T_i is greater than t as it could fall between Y_i and t. So, Zheng and Klein (1994) state that the estimator for $S_T(t)$ that comes from the concept of self-consistency given in Efron (1967) is

$$\hat{S}_T(t) = \frac{1}{n} \left\{ \sum_{i=1}^{n} I(Y_i \geq t) + \sum_{Y_i < t} (1 - \Delta_i) \hat{P}[T > t | T > Y_i, C = Y_i] \right\}. \tag{2.17}$$

where $\Delta_i = I(Y_i = T_i)$. A similar argument can be used when obtaining an estimator of the marginal survival function of C, which gives

$$\hat{S}_C(t) = \frac{1}{n} \left\{ \sum_{i=1}^{n} I(Y_i \geq t) + \sum_{Y_i < t} \Delta_i \hat{P}[C > t | C > Y_i, T = Y_i] \right\}. \tag{2.18}$$

Zheng and Klein (1994) show that when T and C are dependent with copula $\mathcal{C}$, the estimated probabilities in (2.17) and (2.18) can be written as

$$\hat{P}(T > t|T > Y_i, C = Y_i) = \frac{1 - \mathcal{C}_v(1 - \hat{S}_T(t), 1 - \hat{S}_C(Y_i))}{1 - \mathcal{C}_v(1 - \hat{S}_T(Y_i), 1 - \hat{S}_C(Y_i))}$$

and

$$\hat{P}(C > t|C > Y_i, T = Y_i) = \frac{1 - \mathcal{C}_u(1 - \hat{S}_T(Y_i), 1 - \hat{S}_C(t))}{1 - \mathcal{C}_u(1 - \hat{S}_T(Y_i), 1 - \hat{S}_C(Y_i))},$$

where

$$\mathcal{C}_u(a, b) = \frac{\partial \mathcal{C}(u, v)}{\partial u} \quad \text{and} \quad \mathcal{C}_v(a, b) = \frac{\partial \mathcal{C}(u, v)}{\partial v},$$

evaluated at the point $(u, v) = (a, b)$.

Thus the estimators in (2.17) and (2.18) become

$$\hat{S}_T(t) = \frac{1}{n}\left\{\sum_{i=1}^{n} I(Y_i \geq t) + \sum_{Y_i<t}(1 - \delta_i)\frac{1 - \mathcal{C}_v(1 - \hat{S}_T(t), 1 - \hat{S}_C(Y_i))}{1 - \mathcal{C}_v(1 - \hat{S}_T(Y_i), 1 - \hat{S}_C(Y_i))}\right\} \tag{2.19}$$

and

$$\hat{S}_C(t) = \frac{1}{n}\left\{\sum_{i=1}^{n} I(Y_i \geq t) + \sum_{Y_i<t}\delta_i\frac{1 - \mathcal{C}_u(1 - \hat{S}_T(Y_i), 1 - \hat{S}_C(t))}{1 - \mathcal{C}_u(1 - \hat{S}_T(Y_i), 1 - \hat{S}_C(Y_i))}\right\}. \tag{2.20}$$

An iterative process is used to find the self-consistent estimators $\hat{S}_T(t)$ and $\hat{S}_C(t)$. The initial guesses $\hat{S}_T^0(t)$ and $\hat{S}_C^0(t)$ are substituted in the right-hand sides of (2.19) and (2.20) to give $\hat{S}_T^1(t)$ and $\hat{S}_C^1(t)$. This is repeated with $\hat{S}_T^1(t)$ and $\hat{S}_C^1(t)$, and so on until the stable points of the process are found.

This process will converge as the convergence of self-consistent estimators such as those in (2.17) and (2.18) is established in Tsai and Crowley (1985). They found that the EM algorithm can be set up so that it converges to estimators that have the property of self consistency given in Efron (1967), and that convergence is guaranteed as long as the initial estimator used in the algorithm is a step function with mass at the observed time points.

2.9 Copula-Graphic Estimators

Zheng and Klein (1995) suggest an estimator of the marginal survival function of T, $S_T(t)$, based on an assumed copula $\mathcal{C}$, known as the copula-graphic estimator. It is a step function with jumps at the distinct event times. They also define a similar estimator for $S_C(t)$, the marginal survival function of C. This is a step function with jumps at the observed censoring times, which is needed in the estimation of the copula-graphic estimator, $\hat{S}_T(t)$.

We observe $Y_i = \min(T_i, C_i)$ and an indicator function $\Delta_i = I(Y_i = T_i)$ for the ith individual in the data set. The times $t_1, \ldots, t_m$ are the distinct times at which individuals

experience an event or are censored. Let $t_0 = 0$ and $\hat{S}_T(t_0) = \hat{S}_C(t_0) = 1$. If the observation at time t_i is a failure time, then $\hat{S}_C(t_{i-1})$ is used instead of $\hat{S}_C(t_i)$ when computing our estimator, as there will be no change in value for $\hat{S}_C$ at time t_i if the observation is a failure time. Similarly, if at time t_i there is a censored observation, let $\hat{S}_T(t_i) = \hat{S}_T(t_{i-1})$. So if $\Delta_i = 1$, then

$$\hat{S}_T(t_i) + \hat{S}_C(t_{i-1}) - 1 + \mathcal{C}\left[1 - \hat{S}_T(t_i), 1 - \hat{S}_C(t_{i-1})\right] = \frac{1}{n}\sum_{j=1}^{n} I(Y_j > t_i). \tag{2.21}$$

Similarly, if $\Delta_i = 0$, then

$$\hat{S}_T(t_{i-1}) + \hat{S}_C(t_i) - 1 + \mathcal{C}\left[1 - \hat{S}_T(t_{i-1}), 1 - \hat{S}_C(t_i)\right] = \frac{1}{n}\sum_{j=1}^{n} I(Y_j > t_i). \tag{2.22}$$

To ensure that these estimators can handle tied observation times, if there are both failures and censored observations at the time t_i, it is assumed that the censored times occur at time t_i^+ after the failure times. This is a standard assumption when there are ties in the data, and ensures that the estimator $\hat{S}_C(t)$ does not have any jumps at exactly the same time as $\hat{S}_T(t)$. Also any individuals that are censored at t_i would be included in the sum $\sum_{j=1}^{n} I(Y_j > t_i)$ when computing $\hat{S}_T(t_i)$. However, failures that occur at time t_i would not be included in this summation when computing $\hat{S}_C(t_i)$.

2.9.1 Closed form copula graphic estimators

Rivest and Wells (2001) show that in certain circumstances, it is possible to obtain a closed form of copula-graphic estimator of the marginal survival function of T, $\hat{S}_T(t)$. They assume that the joint survival function $\bar{H}(t, c)$ is given by an Archimedean copula, which is a copula of the form given in (2.16). They present the closed form of the estimator when there are no ties in the data. However, we extended their closed-form estimator to data with tied observation times by assuming that if there are both failures and censored observations at time t_i, then the censored observations occur at time t_i^+, just after the failures.

So, if the joint survival function $\bar{H}(t, c)$ can be expressed by an Archimedean copula with generator function $\varphi(t)$, it can be shown that the copula-graphic estimator of the marginal survival function of T is

$$\hat{S}_T(t_i) = \varphi^{-1}\left[-\sum_{Y_i \le t_i, \delta_i = 1}\left\{\varphi\left(\frac{n_i}{n}\right) - \varphi\left(\frac{n_i - d_i}{n}\right)\right\}\right] \tag{2.23}$$

where n_i is the number of individuals at risk at time t_i and d_i is the number of observed failures at time t_i.

2.10 Applying Estimators that use an Assumed Copula to Liver Registration Data Set

In this section, self-consistent and copula graphic estimates for the Liver Registration data set are presented, for different families of copulas. Figure 2.5 shows the self-consistent estimates based on the Gumbel-Hougaard, Frank, gamma frailty and Clayton copulas, each with Kendall's τ of 0.2. Similarly, Figure 2.6 shows the corresponding copula graphic estimates. In Figures 2.5 and 2.6, the Kaplan-Meier estimate of the survival function is also plotted for comparison.

In Figure 2.7 the copula graphic estimates using the Gumbel-Hougaard, Frank and Clayton copulas are presented. However, the estimates here were obtained using the closed forms of the estimators. This form of the estimator gives different results for some of the copulas. This is because Zheng and Klein (1995) use the assumed copula to give the joint distribution function, whereas Rivest and Wells (2001) use the assumed copula to give the joint survival distribution. So if the corresponding survival copula for the Archimedean family had been used, then the two estimators would be the same. As we can see, the closed-form copula graphic estimator for the Clayton copula is the same as the copula graphic estimator for the gamma frailty copula. This is because the gamma frailty copula is the corresponding survival copula for the Clayton copula. For some families of copulas, it does not matter whether the standard copula or the survival copula is used. This is true of the Frank copula as it gives a symmetric distribution to the dependence between the two variables.

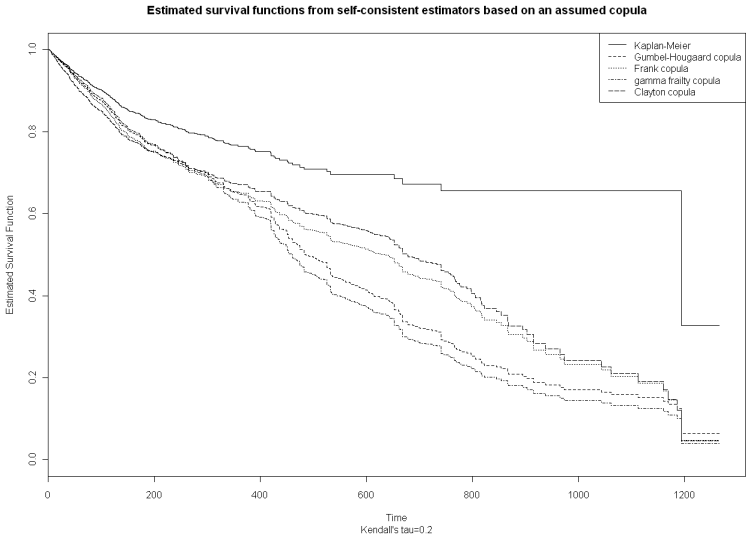

Figure 2.5: Plot of self-consistent estimates based on different assumed copulas for Kendall's $\tau = 0.2$

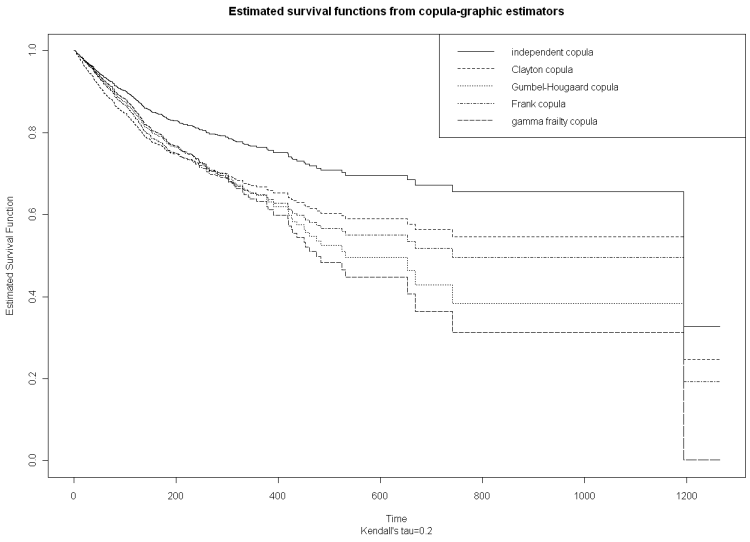

Figure 2.6: Plot of copula-graphic estimates using different assumed copulas for Kendall's $\tau = 0.2$

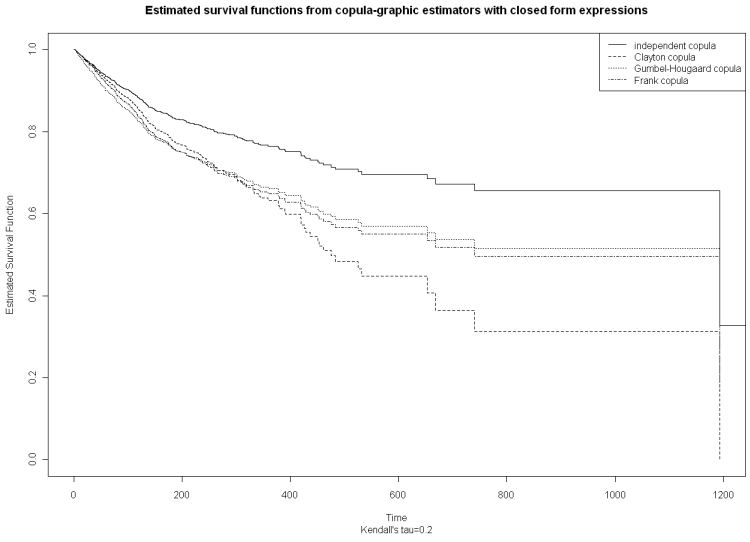

Figure 2.7: Plot of copula-graphic estimates with closed form expression using different assumed copulas for Kendall's $\tau = 0.2$

Here the estimates for a range of copula families for a given value of Kendall's τ have been plotted in each of the Figures 2.5, 2.6 and 2.7. This is so the differences in the estimates given by each family can be seen. However, when applying these methods to data, a copula family that reflects the suspected dependence structure should be chosen. This is because copulas cannot to be fitted to the data to see which fits best as the failure and censoring time are not both observed for each individual.

2.10.1 Selecting copula family

Plots of the joint density functions of the Clayton, Frank, Gumbel-Hougaard and gamma frailty copulas are given in Figure 2.4. These plots should be used to select which of the copulas is believed to be the most appropriate for the data set under consideration. This can be done by choosing the copula with the joint density function that seems most the plausible for our data set.

As the copula $\mathcal{C}$ is used to specify the joint distribution of two variables, then the copula density function $c(u,v) = \frac{\partial^2}{\partial u \partial v}\mathcal{C}(u,v)$ gives the joint density function of the two variables. So a point (u,v) on the surfaces given in the plots in Figure 2.4 corresponds to $f(x,y)$ where $x = S_T^{-1}(u)$ and $y = S_C^{-1}(v)$.

As generally the values of T and C observed tend to be fairly small with only a few individuals having large observations, a copula that has higher density values for low values of u and v should be chosen. From Figure 2.4, it can be seen that a sensible choice would be either the Clayton copula or the gamma frailty copula. As the Clayton copula only gives large density to very small values of u and v, the gamma frailty copula is recommended as its density function does not have such a steep slope.

It is not possible to estimate Kendall's τ so assumed values of this measure are used. The estimators given when using these values of Kendall's τ can be used as bounds for the estimated survival function, if it is believed that the true value of Kendall's τ lies in the assumed interval. Figure 2.8 gives the bounds on the survival function for the Liver Registration data set given by a copula-graphic estimator if Kendall's τ lies between -0.2 and 0.2, assuming firstly a Clayton copula and then assuming a gamma frailty copula.

The plots in Figure 2.8 show how different the bounds on the marginal survival function obtained are when using different copula families. When $t < 200$ the bounds given by the estimators using a gamma frailty copula are tighter than those given by the estimators that use a Clayton copula. Shortly after this time, the bounds given by the estimators that use a gamma frailty copula become much wider than those given by the estimators that use a Clayton copula.

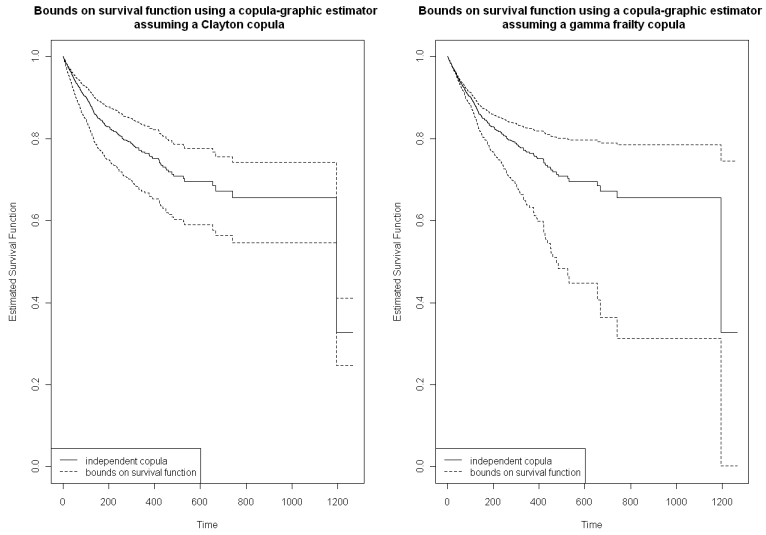

Figure 2.8: Bounds for estimated survival function given by copula-graphic estimators if Kendall's τ lies between -0.2 and 0.2, assuming a Clayton copula and a gamma frailty copula.

2.11 Inclusion of Covariates

As the estimators considered in this chapter are extensions of the Kaplan-Meier estimator, they share some of the limitations of this estimator. The main one here being that not all covariates can be incorporated when using this estimator. In fact, only covariates with simple structures like factors with only a few levels can be considered. For these variables, an estimator for each level of the factor or for each combination of levels of several different factors can be produced. But there is no way of incorporating continuous variables, and some of the most important covariates in the data set under consideration are continuous.

In Yan (2007), it is detailed how it is possible to incorporate covariates into copulas in two ways. Firstly, the margins can be modelled using regression models instead of just a distribution. This means that one of the parameters of the distribution is replaced by $\mathbf{X}^T\boldsymbol{\beta}$, where $\mathbf{X}$ is a vector of covariates and $\boldsymbol{\beta}$ is a vector of parameters. Similarly the copula parameters could be replaced by $\mathbf{X}^T\boldsymbol{\beta}$ to allow covariates to be incorporated.

However, for the estimators considered here that use an assumed copulas, it is the marginal distributions used in the copulas that are being estimated so they cannot be replaced by regression models. Also the value of the copula parameter is used to control the amount of dependence between T and C so the parameter can not be replaced with $\mathbf{X}^T\boldsymbol{\beta}$. So the only way of incorporating covariates in these estimators are those that were discussed previously for the other estimators in this chapter.

2.12 Another estimator of the marginal survival function for failure time variable

There are other estimators that are modifications of the Kaplan-Meier estimator that allow for informative censoring in a data set. One of these estimators is considered briefly, but not applied to the Liver Registration data set.

The estimator is given in Link (1989), uses a frailty model to account for informative censoring. It is assumed that each life time T has a random variable Z associated with it, which is called the frailty. Censoring is then only possible for a subset A of the values of Z. It is usually assumed that censoring is possible for individuals with either high or low frailty.

The survival function $S(t|Z = z)$ is decreasing in Z, so the individuals that tend to have smaller lifetimes are those with high frailties. So if it is assumed that individuals with high frailties are those at risk of censoring and $A = \{z|z \geq a\}$, then there will be heavier censoring on small observations and less censoring on larger observations than under non-informative censoring. This means the Kaplan-Meier estimator would over estimate the survival function $S(t)$ if it was used here.

As before, the observed data are the observation times Y and the indicator function $\Delta = I(Y = T)$. Here the ordered observation times $y_{(i)}$ will be used along with $\Delta_{(i)}$, which is the corresponding value of the indicator function.

The Kaplan-Meier estimator can be written as

$$\hat{S}(t) = \frac{1}{n}\left\{\sum_{i=1}^{n} I(y_{(i)} > t) + \sum_{i=1}^{n}(1 - \Delta_{(i)})P(T > t|Y = y_{(i)}, \Delta = 0)\right\},$$

where

$$P(T > t|Y = y_{(i)}, \Delta = 0) = S(t)/S(y_{(i)}).$$

In the alternative frailty model, proposed in Link (1989), this probability is given by

$$P(T > t|Y = y_{(i)}, |\Delta = 0) = \frac{S(t|Z \in A)}{S(y_{(i)}|Z \in A)}.$$

Then a modified form of the Kaplan-Meier estimator can be obtained by using the algorithm

$$\tilde{S}^{(k+1)}(t) = \frac{1}{n}\left\{\sum_{i=1}^{n} I(y_{(i)} > t) + \sum_{i=1}^{n}(1 - \Delta_{(i)})\frac{\tilde{S}^{(k)}(t|Z \in A)}{\tilde{S}^{(k)}(y_{(i)}|Z \in A)}\right\},$$

and letting $k \to \infty$.

This estimator is not applied to the data set under consideration for two reasons. Firstly, as it just another extension of the Kaplan-Meier estimator then it still has all the disadvantages associated with this estimator and so does not provide an improvement

on any of the other estimators considered here. Secondly, and more importantly, the assumption used here is not a realistic one for the situation being considered. For this method, it is assumed that censoring is possible for only a subset of the individuals in the data set. Although there are patients on the waiting list for a liver transplant that have a greater hazard of receiving a transplant, we do not want to restrict censoring to just these individuals. This would imply that there are patients on the waiting list who could not receive a transplant and this is not realistic.

2.13 Discussion

The bounds that are given by the all estimators applied to the Liver Registration data set in this chapter are not useful in a practical setting as they are too wide. The method used in both Slud and Rubinstein (1983) and Klein and Moeschberger (1988) of assuming that the parameter controlling the dependence lay within a restricted region gave bounds that were tighter than those of Peterson (1976). However, even these were still too wide to be of use. It is still possible to use the estimators with an assumed value of dependence. But they still not of much use in a practical application as they do not allow all the important covariates to be incorporated.

A number of different estimators that extend the Kaplan-Meier estimator to allow for informative censoring have been presented here. Some of the estimators are preferable to others. In particular, it is not as easy to specify an interpretable amount of dependence for the Fisher-Kanarek and Slud-Rubinstein estimators as for the other estimators that use Kendall's τ. However, the Fisher-Kanarek estimator does allow for non-informative censoring as well as informative censoring, unlike the rest of the estimators which only consider one type of censoring. This means either only the data up until the first non-informative censoring is used or the non-informative censoring is treated as informative censoring. Here the latter method is used.

However, use of the Fisher-Kanarek estimator is still not recommended. This is because when the last observation in a data set is censored we can see some strange behaviour in the estimate of the survival function when assuming a positive dependence between T and C. The reason for this is discussed in detail in Section 2.6.

Zheng and Klein (1994) present the results of a small simulation study that compares the self consistent estimator and the copula-graphic estimator. A gamma frailty copula with exponential margins is used with a value of θ that gives $\tau = 0.5$. The parameters of the exponential margins chosen give $P(X < Y) = 0.50$, which equates to 50% censoring. Also a sample size of 20 is used. They calculate the relative biases of the estimators as the marginal distribution function $F_T(t)$ increases. The relative bias of $\hat{S}$ is defined

as $(E[\hat{S}(t_p)] - p)/p$ at time t_p where $S(t_p) = p$. Both estimators are biased for large t, as the size of the risk set becomes smaller. However, the self-consistent estimator has a significantly larger bias than the copula-graphic estimator. This behaviour is typical for these estimators as they also carried out further simulation studies for other copulas, association parameters and sample sizes. The results of these simulation studies are given in Zheng (1992).

Also the copula-graphic estimator is less computationally intensive as it only requires one pass through the data to construct the estimator. In contrast the self-consistent estimator requires a pass through the data at each iteration. For these reasons the copula-graphic estimator is recommended rather than the self-consistent estimator.

Although the preferred estimator that uses an assumed copula is now known, it is still not known how it compares to the other estimators in this chapter. Also, it is not known whether the gamma frailty copula that is recommended here gives estimators that are closer to the true survival function than the other copulas. However, as these methods cannot easily be used in practice due to the wide bounds found and the difficulties with incorporating covariates, it would not be particularly useful to identify the preferred estimator of those detailed in this chapter. Therefore, in the following chapter, we go on to look at more recent approaches to account for informative censoring that allow the use of a wider variety of covariates.

Chapter 3

Estimation when using Regression Models for the Censoring Process and Sensitivity Analyses under Informative Censoring

This chapter continues the literature review that was started in Chapter 2. The estimators considered previously gave bounds on the estimated survival function but it was found that these bounds were too wide to be of much use. These estimators also generally did not allow covariates to included easily.

The methods reviewed in this chapter are of more use in practice than those in Chapter 2 and generally can easily include covariates. They can be split into two categories, estimators that use regression models for the censoring process and sensitivity analyses.

The estimators that use regression models for the censoring process are described in sections 3.1 and 3.2. These include one of the most popular methods in the literature on informative censoring. These are the inverse probability of censoring weighted estimators that are given in Section 3.1.

The sensitivity analyses in sections 3.3 to 3.6 assess the sensitivity of the results from standard models to the assumption of informative censoring. Sensitivity analyses for both standard parametric survival models and Cox's proportional hazards model are presented. In Section 3.7, a sensitivity analysis for an estimator that already accounts for informative censoring is described, that allows us to assess how biased this estimator could be if there is dependence between T and C that is not explained by its assumed dependence structure.

As this is a review chapter, all the methods discussed can be found in the literature and, unless otherwise stated, the original work is the application of the methods to the

Liver Registration data set. Some of the notation used in this chapter may differ from that in the papers referenced as we present all the methods in consistent notation.

3.1 IPCW estimators

Inverse probability of censoring weighted (IPCW) estimators were first introduced by Robins and Rotnitzky (1992) and Robins (1993). They have been recognised as a way to adjust for the bias introduced by dependent censoring, for cases where the same prognostic factors predict both time to failure and time to censoring.

This method relies on the assumption of no unmeasured confounders for censoring or the assumption of sequential ignorability of censoring, which states that if the cause specific hazard of censoring is conditioned on the recorded history,$\bar{\mathbf{V}}(t)$, of a vector of possibly time-dependent covariates,$\mathbf{V}$, then it does not further depend on T,

$$h_C(t|\bar{\mathbf{V}}(t), T, T > t) = h_C(t|\bar{\mathbf{V}}(t), T > t) \tag{3.1}$$

where $\bar{\mathbf{V}}(t)$ is defined as $\{\mathbf{V}(x); 0 \leq x \leq t\}$. If all the prognostic factors are recorded in $\mathbf{V}(t)$ then the IPCW estimators outlined below will adjust completely for the bias due to dependent censoring. However, in practice, we will not be able to record all possible prognostic factors, but if the most important factors are recorded then the use of IPCW estimators will considerably reduce the bias caused by dependent censoring.

Another concept that is necessary to introduce is that of the data being coarsened at random (CAR). This was introduced by Heitjan and Rubin (1991) as a generalisation of the concept of missing at random. Censoring is just one example of how a dataset could be coarsened. Censored data are CAR if the censoring mechanism does not depend on the values of the outcome, although it is allowed to depend on the values of any covariates. The CAR assumption can be expressed as

$$h_C(t|\bar{\mathbf{V}}(T), T, T > t) = h_C(t|\bar{\mathbf{V}}(t), T > t). \tag{3.2}$$

This is similar to Equation 3.1, except that $\bar{\mathbf{V}}(t)$ has been replaced by $\bar{\mathbf{V}}(T)$. So CAR implies (3.1), but (3.1) does not imply CAR.

3.1.1 Constructing IPCW estimators

The IPCW versions of the Kaplan-Meier (KM) estimator and the Cox partial likelihood score function are given here. The construction of these types of estimators was outlined in Robins and Finkelstein (2000), who used it to adjust for dependent censoring when comparing two treatments in an AIDS clinical trial. IPCW estimators can be found by weighting the contribution of each subject by the inverse of an estimate of the conditional

probability of having remained uncensored until time t. The effect that this weighting has on estimators will now be explained, using the KM estimator as an example. The standard KM estimator is given by

$$\hat{S}(t) = \prod_{\{i;Y_i<t\}} \left(1 - \frac{d_i}{n_i}\right),$$

where d_i is the number who fail at time Y_i and n_i is the number at risk at time Y_i. For the IPCW KM estimator, the contributions to these two terms are weighted by the inverse probability of remaining uncensored until time t. Then the numerator of the fraction estimates the number of individuals who would have been observed to fail at time Y_i in the absence of any censoring. Similarly, the denominator estimates the number of subjects at risk at time Y_i in the absence of any censoring. This means that the IPCW KM estimator gives an estimate of the survival function in the absence of censoring.

To construct the weights an estimate of the probability of remaining uncensored until time t given $(\bar{\mathbf{V}}(T), T)$ is needed, where $\bar{\mathbf{V}}(T)$ is the recorded history of a covariate vector up until time T. This is given by a KM estimator for time to censoring that has been extended to include time-dependent covariates. The model for censoring that will be used is

$$h_C(t|\bar{\mathbf{V}}(t), T > t) = h_0(t) \exp\{\boldsymbol{\beta}_C' \mathbf{V}(t)\}, \tag{3.3}$$

where $h_0(t)$ is the baseline hazard and $\boldsymbol{\beta}_C$ is a vector of parameters. A Cox proportional hazards model will be fitted to give the partial likelihood estimate $\hat{\boldsymbol{\beta}}_C$. The observed values are denoted by $Y = \min(T, C)$. Indicator values $R(u) = I(Y \geq u)$ and $\Delta = I(T = Y)$ are used to identify those at risk and which observations are failures. Let $t_1, t_2, \ldots, t_n$ be the times of the observations.

In the literature, several different ways of selecting the covariates to be included in $\mathbf{V}(t)$ have been suggested. We consider the ways suggested in Robins and Finkelstein (2000), Schaubel et al. (2009) and Zhang and Schaubel (2010). These are outlined in more detail in Section 3.1.3. When we apply the method to our data set, we shall compare the IPCW estimates given by using each of these ways.

Under CAR and model (3.3), then it is possible to derive the following KM estimator for censoring

$$\hat{K}_i^{\mathbf{V}}(t) = \prod_{\{j;t_j<t,\Delta_j=0\}} [1 - \hat{h}_0(t_j) \exp\{\hat{\boldsymbol{\beta}}_C' \mathbf{V}_i(t_j)\}], \tag{3.4}$$

where the Cox estimator of the baseline hazard for censoring at observation time t_j is given by

$$\hat{h}_0(t_j) = \frac{(1 - \Delta_j)}{\{\sum_{i=1}^n \exp(\boldsymbol{\beta}_C' \mathbf{V}_i(t_j)) R_i(t_j)\}}.$$

The estimator in (3.4) is written as $\hat{K}_i^{\mathbf{V}}(t)$ to show that it depends on $\mathbf{V}_i(t)$. The usual KM estimator of the probability of being uncensored at time t is denoted by $\hat{K}_i^0(t)$. This will be equal to $\hat{K}_i^{\mathbf{V}}(t)$ when $\boldsymbol{\beta}_C$ is the zero vector.

We can define subject specific weights, $\hat{W}_i(t)$, which will be used in the IPCW versions of the KM estimator and Cox partial likelihood. One possible weight is $\hat{K}_i^0(t)/\hat{K}_i^{\mathbf{V}}(t)$ which will be close to one for all t if and only if $\bar{\mathbf{V}}(t)$ does not predict the hazard of censoring at t. So, if we do have informative censoring, $\hat{W}_i(t)$ will not be close to one. Another weight that could be used is $1/\hat{K}_i^{\mathbf{V}}(t)$. However, as shown in Robins (1993), using $\hat{K}_i^0(t)/\hat{K}_i^{\mathbf{V}}(t)$ as the weight has important efficiency advantages. These will be discussed further in Section 3.1.2. They suggest that using $1/\hat{K}^{\mathbf{V}}(t)$ as the weight may be appropriate when there is only light or moderate censoring but if there is heavy censoring, this value can become quite large. Because of this, Robins and Finkelstein (2000) recommend using $\hat{K}_i^0(t)/\hat{K}_i^{\mathbf{V}}(t)$ when there is heavy censoring. From now on, $\hat{K}_i^0(t)/\hat{K}_i^{\mathbf{V}}(t)$ will be referred to as a "stabilised" weight and $1/\hat{K}_i^{\mathbf{V}}(t)$ as an "unstabilised" weight.

Now it is possible to define the IPCW Kaplan-Meier estimator for time to failure. It is shown in Robins and Finkelstein (2000) that the value of this estimate at time t is

$$\hat{S}_T(t) = \prod_{\{i;t_i<t\}} \left(1 - \frac{\Delta_i \hat{W}_i(t_i)}{\left\{\sum_{k=1}^n R_k(t_i)\hat{W}_k(t_i)\right\}}\right). \tag{3.5}$$

It does not matter whether $\hat{K}_i^0(t)/\hat{K}_i^{\mathbf{V}}(t)$ or $1/\hat{K}^{\mathbf{V}}(t)$ is used for $\hat{W}_i$ in (3.5) as $\hat{K}^0(t)$ cancels from both the numerator and the denominator. Therefore the merits of using stabilised weights instead of unstabilised weights only need to be considered when using the Cox partial likelihood.

The IPCW Cox partial likelihood score for a vector of parameters $\boldsymbol{\beta}_T$, is also derived in Robins and Finkelstein (2000) and is given by

$$U(\boldsymbol{\beta}_T) = \sum_i \Delta_i \hat{W}_i(t_i) \left[\mathbf{Z_i} - \frac{\sum_{j=1}^n R_j(t_i)\hat{W}_j(t_i)\mathbf{Z_j}e^{\boldsymbol{\beta}_T'\mathbf{Z_j}}}{\sum_{j=1}^n R_j(t_i)\hat{W}_j(t_i)e^{\boldsymbol{\beta}_T'\mathbf{Z_j}}}\right] \tag{3.6}$$

where $\mathbf{Z}$ is a vector of baseline covariates to be included in the model for time to failure. When fitting a model with weights like this, robust estimates of the variance of the parameter estimates need to be used.

Weibull model for time to censoring Robins and Finkelstein (2000) only consider a Cox model for time to censoring, but it is also possible to use a Weibull model for the baseline hazard for censoring, where

$$h_{C0}(t) = \lambda\eta t^{\eta-1},$$

and the survival function for time to censoring for individual i can be estimated by

$$\hat{S}_{Ci}(t) = \exp\left\{-\exp(\hat{\boldsymbol{\beta}}'_W \mathbf{V}_i(t))\hat{\lambda}t^{\hat{\eta}}\right\} \tag{3.7}$$

where $\hat{\boldsymbol{\beta}}_W$ is obtained using the Weibull proportional hazards model. However, this does mean that it is not easy to include time-dependent covariates in the model. So, when using these weights to obtain IPCW estimates for the Liver Registration data set, we will consider only time independent covariates when using a Weibull proportional hazards model.

Again it is possible to use both "stabilised" and "unstabilised" weights to obtain IPCW estimates. The unstabilsed weights are $1/\hat{S}^{\mathbf{V}}_{Ci}(t)$ and are comparable to $1/\hat{K}^{\mathbf{V}}_i(t)$ used previously. The stabilised weights require the use of $\hat{S}^0_{Ci}(t)$, which is the estimated survival function in (3.7) with $\mathbf{V}_i(t)$ replaced by the zero vector. The weight used is then $\hat{S}^0_{Ci}(t)/\hat{S}^{\mathbf{V}}_{Ci}(t)$.

We expect that the IPCW estimates using a Cox model for censoring and a Weibull model for censoring will be similar for the Liver Registration data set. This is because we have little information on how the UKELD score changes over time in this data set. However, if more information on time-dependent covariates is available then it is recommended that the Cox model for censoring is used, because it can easily incorporate time-dependent covariates.

3.1.2 Stabilised weights vs. unstabilised weights

In this section, we will discuss whether the stabilised weights $\hat{K}^0_i(t)/\hat{K}^{\mathbf{V}}_i(t)$ or the unstabilised weights $1/\hat{K}^{\mathbf{V}}_i(t)$ should be used when calculating IPCW estimates. The weights being considered here are those that use Cox models for censoring as these are the weights used in Robins (1993a), which established many of the results on the properties of the estimators that are presented in this section.

Robins and Finkelstein (2000) recommend using $\hat{K}^0_i(t)/\hat{K}^{\mathbf{V}}_i(t)$ as the subject specific weight as this gives important efficiency advantages. In this context, the estimate with the lowest variance is regarded as the most efficient. Semi-parametric variance bounds for the semi-parametric models detailed in Section 3.1.1 are given in Robins and Rotnitzky (1992) and Robins (1993b). These papers rely heavily on the theory of semi-parametric efficiency bounds given in Newey (1990) and Bickel at al. (1998)[1].

Robins (1993a) proves that the solution $\hat{\boldsymbol{\beta}}_T$ to $U(\boldsymbol{\beta}_T) = 0$ is consistent and asymptotically normal when $\hat{K}^0_i(t)/\hat{K}^{\mathbf{V}}_i(t)$ is used for $\hat{W}(t)$ in (3.6), given that (3.1) holds and the model for time to censoring is correctly specified. The solution $\hat{\boldsymbol{\beta}}_T$ to $U(\boldsymbol{\beta}_T) = 0$ remains

[1]The publication date of this book is after the publication of the papers by Robins (1993) and Robins and Rotnitzky (1992) but both papers include an advance manuscript in their bibliographies.

consistent and asymptotically normal if $1/\hat{K}_i^{\mathbf{V}}(t)$ is used for $\hat{W}(t)$ in (3.6). However, if $\hat{K}_i^0(t)/\hat{K}_i^{\mathbf{V}}(t)$ is used then the estimator $\hat{\boldsymbol{\beta}}_T$ is asymptotically more efficient than the usual Cox partial likelihood estimator of $\boldsymbol{\beta}_T$, if

$$h_C(t|\bar{\mathbf{V}}(t), T > t) = h_C(t, T > t) \tag{3.8}$$

and (3.1) hold, that is if there is non-informative censoring. This result suggests that we should use an IPCW estimator with stabilised weights instead of the Cox partial likelihood estimator of β_T, even when the censoring is non-informative. However, many of the results given in this section rely on the correct specification of the Cox model for time to censoring. We cannot be sure that the model for time to censoring used is correct and therefore use of the usual Cox partial likelihood estimator of β_T is justified.

As in the liver transplantation setting being considered in this thesis it is unlikely that (3.8) always holds, then these efficiency advantages are not as important as choosing weights that reflect the situation under consideration. There will be some individuals on the waiting list for a liver transplant who are at much greater risk of being censored than others, so these individuals would need to be more heavily weighted. Therefore, we recommend that the unstabilised weights should be used in the liver transplant setting rather than the stabilised weights.

3.1.3 Models for censoring process

There are several models for time to censoring that have been suggested in the literature. The first model for time to censoring considered here is suggested by Robins and Finkelstein (2000), where only the time-dependent covariates that are significant for both time to failure and time to censoring are included. As the assumption of sequential ignorability of censoring relies on all the shared prognostic factors being included in the model for time to censoring, we recommend that this model is used unless there is a good argument for using one of the following models.

The next model used for time to censoring includes all the baseline variables that are to be included in the time to failure model plus time-dependent UKELD. This model was proposed in Schaubel et al. (2009). The final model used includes any baseline covariates that were found to be significant for time to censoring plus time-dependent UKELD. Use of such a model was suggested in Zhang and Schaubel (2010).

As all these use Cox models for time to censoring, we can define the models considered here as

Cox model 1 which uses just time-dependent UKELD,

Cox model 2 which uses primary liver disease category, ethnicity, age, serum sodium at time of registration, INR at time of registration and time-dependent UKELD, and

Cox model 3 which uses primary liver disease category, ethnicity, age, serum sodium at time of registration, INR at time of registration, height, blood group and time-dependent UKELD.

The baseline covariates that were found to be significant for time to censoring were all the variables in the model for time to failure plus two additional covariates, so we are successively adding more covariates in the models considered above.

Weibull models for time to censoring are also considered. So that they are comparable to the models defined above, the same covariates will be used, except that time-dependent UKELD will be replaced by the value of UKELD at the time of registration. So we define these models as

Weibull model 1 which uses just UKELD score at time of registration,

Weibull model 2 which uses primary liver disease category, ethnicity, age, UKELD score at time of registration, serum sodium at time of registration and INR at time of registration, and

Weibull model 3 which uses primary liver disease category, ethnicity, age, UKELD score at time of registration, serum sodium at time of registration, INR at time of registration, height and blood group.

3.1.4 Application to the Liver Registration data set

Firstly, IPCW KM estimators using each of the models described in Section 3.1.3 are fitted to the Liver Registration data set. Figures 3.1, 3.2 and 3.3, compare the IPCW KM estimators using Cox and Weibull models for censoring to the standard KM estimator of the marginal survival function. We see that all the plots in Figures 3.1, 3.2 and 3.3 give similar IPCW KM estimators that do not deviate greatly from the standard KM estimator. This suggests that the potentially informative censoring in the Liver Registration data set has little effect on the estimate of the survival function. This does not agree with the estimates of the survival function found in Chapter 2, which suggested that even a small amount of dependence between T and C would result in a fairly large change in the estimate of the survival function.

One possible reason why the IPCW KM estimator does not vary greatly from the standard KM estimator is that the dependence between T and C is not completely due

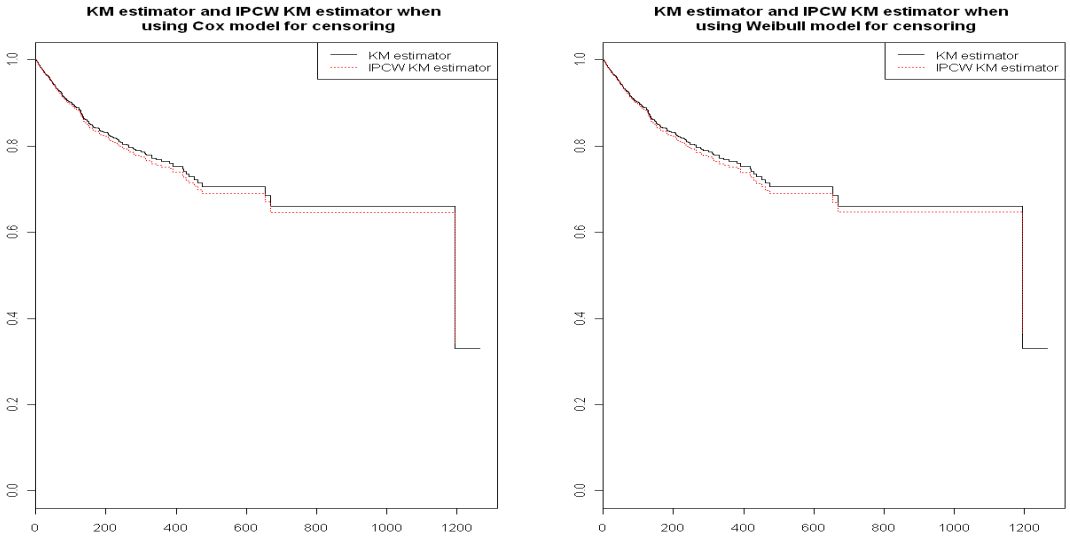

Figure 3.1: Plots comparing IPCW KM estimators with unweighted KM estimators, using Cox Model 1 and Weibull Model 1 for censoring respectively

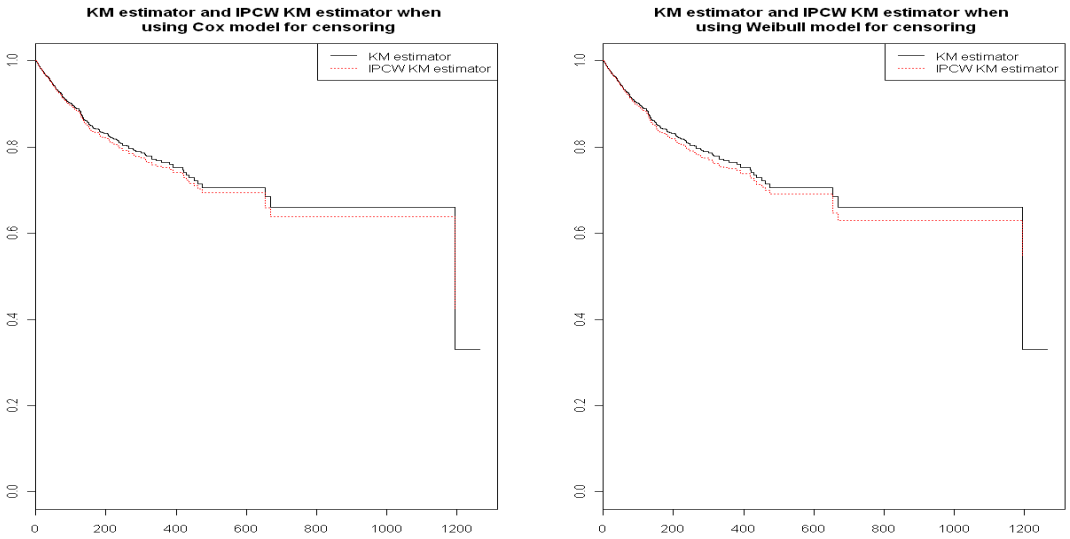

Figure 3.2: Plots comparing IPCW KM estimators with unweighted KM estimators, using Cox Model 2 and Weibull Model 2 for censoring respectively

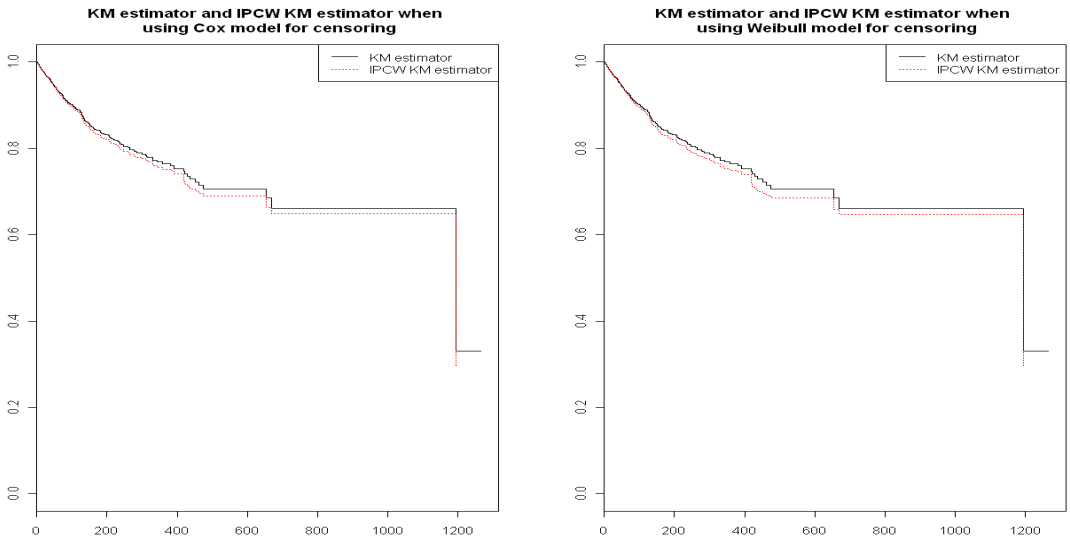

Figure 3.3: Plots comparing IPCW KM estimators with unweighted KM estimators, using Cox Model 3 and Weibull Model 3 for censoring respectively

to shared prognostic factors included in the model for time to censoring. There could be residual dependence caused by unmeasured prognostic factors. Scharfstein and Robins (2002) and Rotnitzky et al. (2007) developed methods that allow the effect of residual dependence on an estimator that assumes sequential ignorability of censoring to be assessed. This is covered in more detail in Section 3.7. Unfortunately, the estimator considered in Section 3.7 is not the IPCW KM estimator presented in Section 3.1.1, so the effect of possible residual dependence on the IPCW KM estimate of the survival function cannot be assessed.

However, this analysis using the IPCW KM estimator is fairly simplistic and does not allow for adjustment for significant covariates for time to failure. Therefore we fit IPCW Cox models for time to failure to the Liver Registration data set. These allow us to assess the effect of informative censoring on individual parameter estimates and also the estimated survival function for individuals in the data set.

Several IPCW Cox models for time to death are fitted to the data set. The same baseline covariates will be included in all the models for time for failure. These are primary liver disease category, ethnicity, age, UKELD score at time of registration, serum sodium at time of registration and INR at time of registration. However, different models are used for to time to censoring and the corresponding IPCW estimates for each model are presented, along with the unweighted estimates obtained by fitting the standard Cox model. The models for time to censoring that are used were discussed in Section 3.1.3.

Figures 3.4, 3.5, 3.6 and 3.7 give the point estimates and 95% confidence intervals

obtained by fitting all these models using Cox models for censoring and Weibull models for censoring using both stabilised and unstabilised weights.

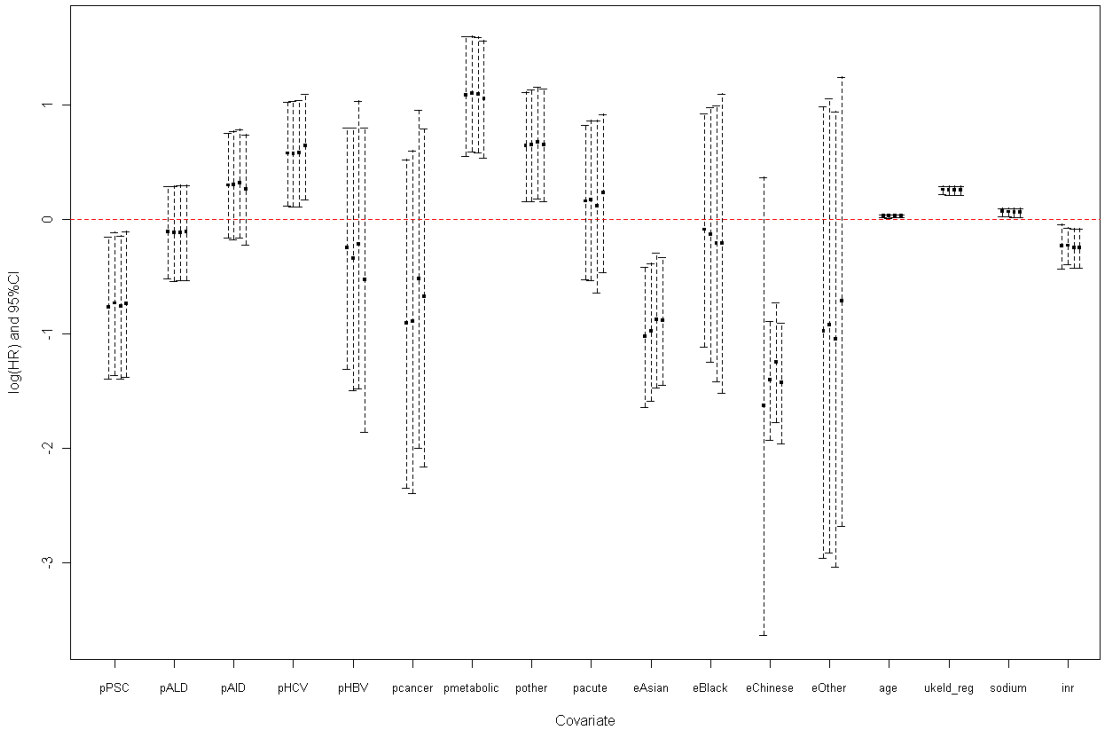

Figure 3.4: Point estimates and 95% confidence intervals for parameters in time to failure model, for unweighted Cox model and IPCW Cox model using Cox models 1, 2 and 3 for time to censoring respectively. All the weights used in IPCW estimates are stabilised.

We find that the IPCW estimates using stabilised weights, which are shown in Figure 3.4, the point estimates are slightly different from the standard point estimates, but generally significant covariates do not become non-significant or vice versa. This is with the exception of some of the estimates for the Chinese level of ethnicity. Under the standard Cox model, this parameter estimate has wide bounds as there are only a small number of individuals with this ethnicity in the data set. However the use of weights here is analogous to the use of sampling weights. This means that the number of observations with this ethnicity is being increased so there is less uncertainty about this parameter estimate.

However, we see that for the IPCW estimates that use unstabilised weights, which can be seen in Figure 3.5, there are more changes from the standard estimates. Several different levels of the categorical variables that are significant under the standard model, become non-significant. However, these changes are likely to be caused by the heavy censoring in the data set making some of these unstabilised weights quite large.

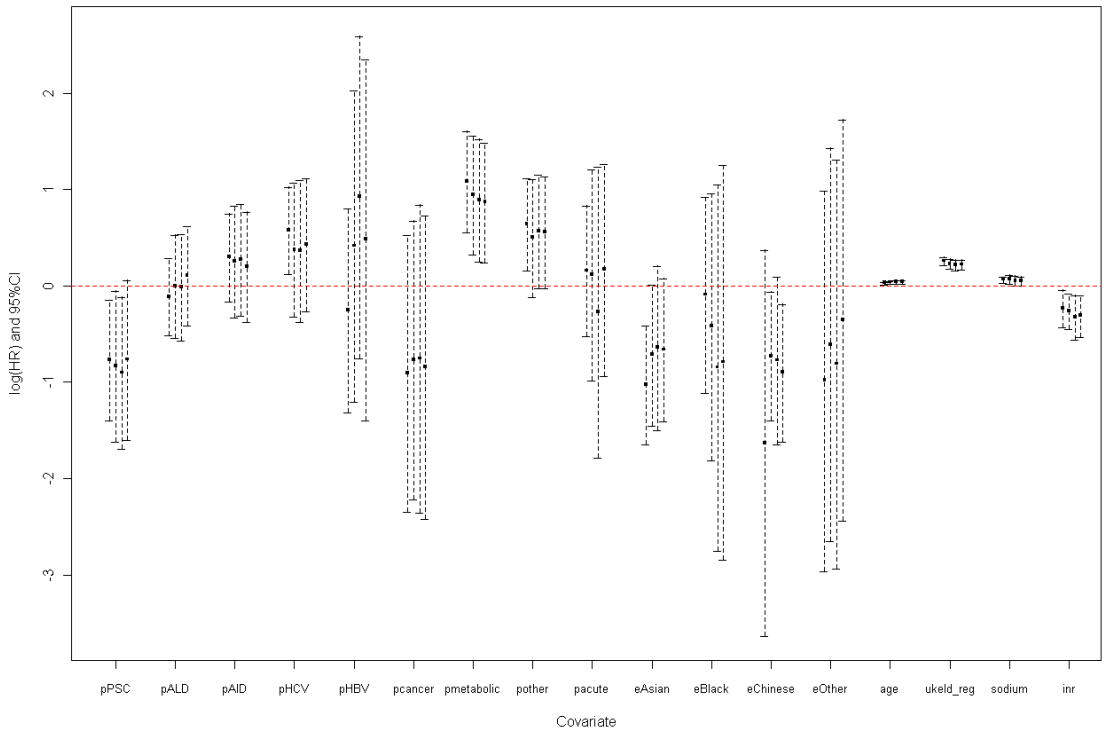

Figure 3.5: Point estimates and 95% confidence intervals for parameters in time to failure model, for unweighted Cox model and IPCW Cox model using Cox models 1, 2 and 3 for time to censoring respectively. All the weights used in IPCW estimates are unstabilised.

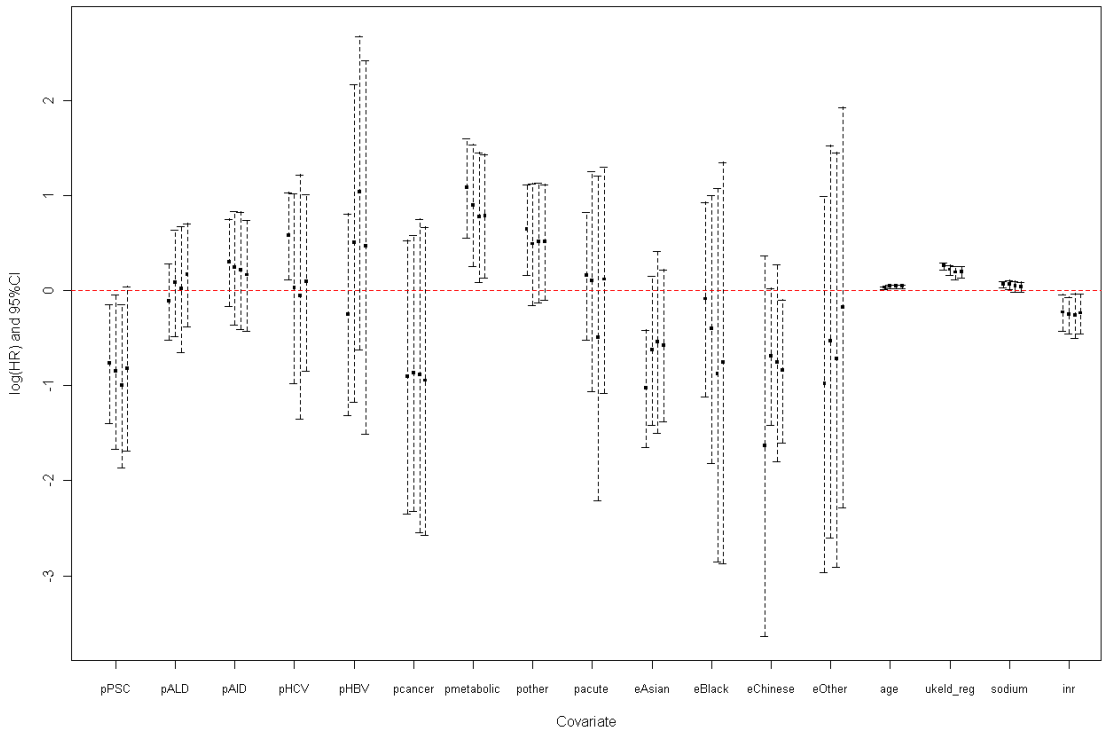

Figure 3.6: Point estimates and 95% confidence intervals for parameters in time to failure model, for unweighted Cox model and IPCW Cox model using Weibull models 1, 2 and 3 for time to censoring respectively. All the weights used in IPCW estimates are stabilised.

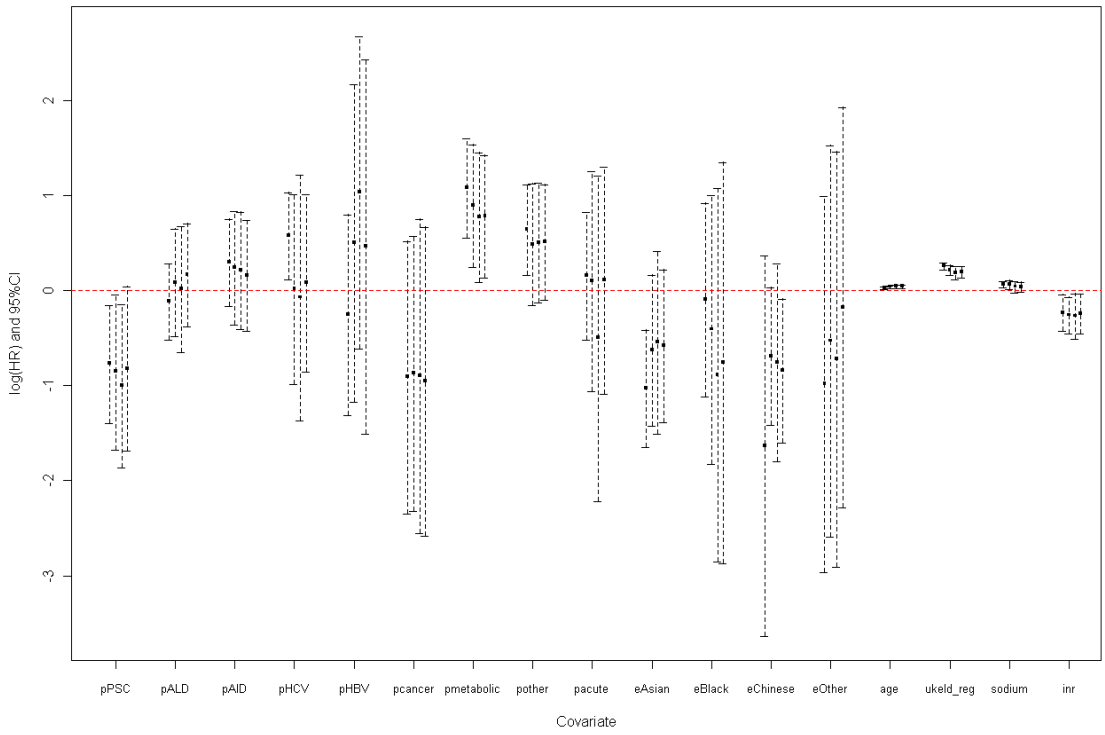

Figure 3.7: Point estimates and 95% confidence intervals for parameters in time to failure model, for unweighted Cox model and IPCW Cox model using Weibull models 1, 2 and 3 for time to censoring respectively. All the weights used in IPCW estimates are unstabilised.

For the covariates and factor levels that remain significant when using an IPCW Cox model, we will examine the changes the estimated hazard ratios. There is a slight decrease in the point estimates of the hazard ratio for patients with metabolic liver disease when using an IPCW Cox model. This suggests that the standard Cox model slightly overestimates the hazard ratio for these patients. The hazard ratios for age, UKELD score, serum sodium and INR all also remain significant when using an IPCW Cox model. However, there is very little difference between the point estimates from the standard Cox model and the point estimates from the IPCW Cox models.

The IPCW estimates using Weibull models for time to censoring, with both stabilised and unstabilised weights, can be seen in Figures 3.6 and 3.7 respectively. The results in these two figures are very similar, suggesting that when using a Weibull proportionals hazards model for time to censoring it does not matter whether stabilised or unstabilised weights are used. The changes from the standard estimates are also similar to those observed in Figure 3.5, with some levels of categorical variables that were significant becoming non-significant.

Again, we examine the changes in the estimated hazard ratios for the covariates and factor levels that remain significant when using an IPCW Cox model. The results are very similar to those for the IPCW estimates given in Figure 3.5. There is slight decrease in the estimated hazard ratio for patients with metabolic liver disease, suggesting the standard Cox model slightly overestimates the hazard ratio for these patients. The hazard ratios for age, UKELD score, serum sodium and INR remain significant, with the exception of a couple of the estimated hazard ratios for serum sodium. Again there is very little change in the point estimates for these covariates.

Figures 3.4 to 3.7 show the effects of inverse probability of censoring weighting on the parameter estimates of the Cox model. We will now look at the effects that these changes in the parameter estimates can have on the survival functions for individuals in the data set.

Figure 3.8 compares the estimated survival function under the standard Cox model with the estimated survival function under the IPCW Cox model for the individual who had the largest observed value of $\hat{\beta}_T^{IPCW'}\mathbf{x}_i - \hat{\beta}_T^{0'}\mathbf{x}_i$. The weights used for the IPCW estimates were unstabilised weights using Cox model 1 for time to censoring. We can see that there is a large difference between the two estimated survival functions. The estimated survival function under the standard model, shown by the solid line in Figure 3.8 does not fall below 0.9, whereas the estimated survival function under the IPCW Cox model, shown by the dashed line, has a median survival time of approximately 1200 days.

The analyses carried out in this section show that using an IPCW version of the KM estimate of the survival function has little effect on the value of the estimated survival

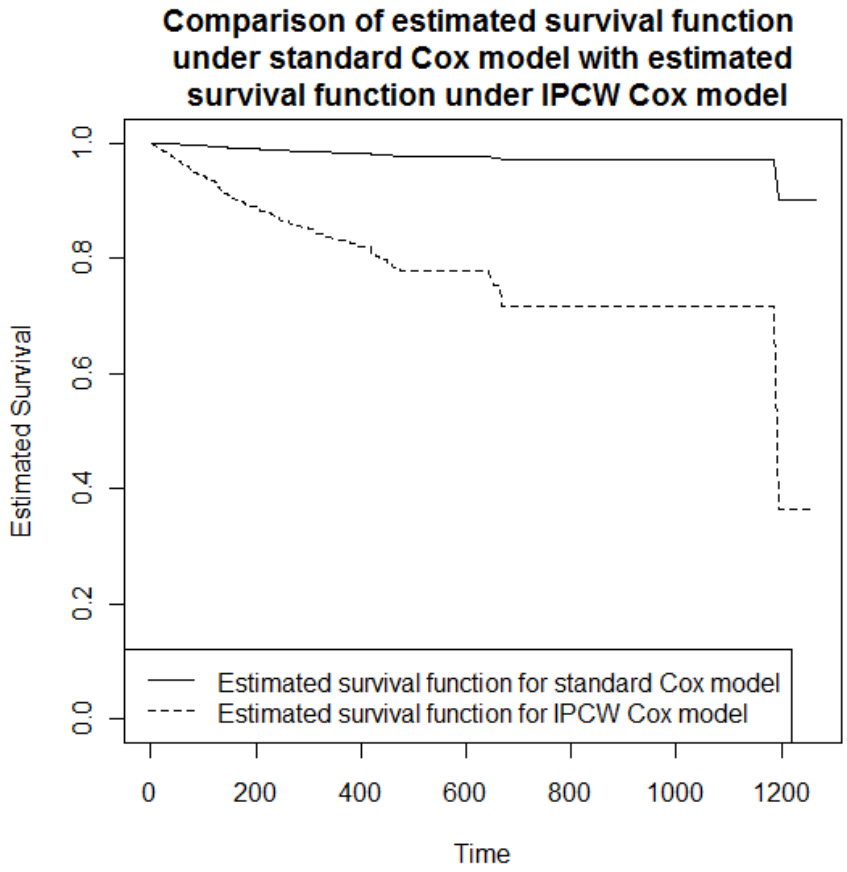

Figure 3.8: Plot comparing the estimated survival function under the standard Cox model with the estimated survival function under the IPCW Cox model for the individual that has the largest observed value of $\hat{\beta}_T^{IPCW'}\mathbf{x}_i - \hat{\beta}_T^{0'}\mathbf{x}_i$. The weights used are unstabilised weights using Cox model 1 for time to censoring.

function. However, if an IPCW Cox model is used, which allows for adjustment for significant covariates, then there can be a large effect on the estimated survival function for some individuals in the data set.

3.1.5 Other weighted estimators

A similar weighted KM estimator is derived in Satten et al. (2001), but using Aalen's additive hazard model instead of the proportional hazards model when calculating $\hat{K}_i^{\mathbf{V}}(t)$. Aalen's model is more flexible than the proportional hazards model as the regression coefficients for the p covariates, $\beta_{C1}(t), \ldots, \beta_{Cp}(t)$, are allowed to change continuously over time. So, Aalen's model has hazard function

$$h_C(t|\bar{\mathbf{V}}_i(t)) = \sum_{k=1}^{p} \beta_{Ck}(t) V_{ik}(t),$$

where $V_{ik}(t)$ is the value of the kth covariate for the ith individual at time t and $V_{i0}(t) = 1$. Also the cumulative hazard function of Aalen's model can be written as

$$\begin{aligned} H(t|\bar{\mathbf{V}}_i(t)) &= \int_0^t h_C(u|\bar{\mathbf{V}}_i(u))du \\ &= \sum_{k=0}^{p} V_{ik}(t) \int_0^t \beta_{Ck}(u)du \\ &= \sum_{k=0}^{p} V_{ik}(t) B_{Ck}(t), \end{aligned} \tag{3.9}$$

where $B_{Ck}(t)$ is the cumulative regression coefficient for the kth covariate. It is easier to estimate the cumulative regression coefficients than the regression coefficients, and Aalen (1989) uses a least-squares-like estimator of $\mathbf{B}_C(t) = (B_{C0}(t), \ldots, B_{Cp}(t))$,

$$\hat{\mathbf{B}}_C(t) = \sum_{i=1}^{n} I(Y_i \leq t)(1 - \Delta_i)\mathbf{A}^{-1}(Y_i)\mathbf{V}_i(Y_i) \tag{3.10}$$

where

$$\mathbf{A}(t) = \sum_{i=1}^{n} I(Y_i \geq t)\mathbf{V}_i(t)\mathbf{V}_i'(t).$$

The estimator in (3.10) can be substituted into (3.9), so that an estimator for $H_C(t|\bar{\mathbf{V}}_i(t))$ is

$$\hat{H}_C(t|\bar{\mathbf{V}}_i(t)) = \sum_{j=1}^{n} I(Y_j \leq t)(1 - \Delta_j)\mathbf{V}_i(Y_j)\mathbf{A}^{-1}(Y_j)\mathbf{V}_j(Y_j), \quad t \leq Y_i,$$

and this can be used to obtain an estimate for $K_i^{\mathbf{V}}(t)$ as

$$\hat{K}_i^{\mathbf{V}}(t) = \prod_{s \leq t} [1 - dH_C(s|\bar{\mathbf{V}}_i(s))].$$

There are two possible problems that can arise when using Aalen's additive hazard model, which is why it is not generally used in practical situations. The hazard estimates may be negative and the estimator involves inverting a matrix that may not have full rank. However, Satten et al. (2001) show that neither of these problems occur when estimating $K_i^{\mathbf{V}}(t)$.

3.2 Other estimators that use models for the censoring process

There are several other estimators that use models for the censoring process when there is informative censoring in a data set. Wu and Carroll (1988) use a linear random effects model, Wu and Bailey (1989) use a conditional linear model and Schlucter (1992) uses a log-normal survival model that is an extension of the linear random-effects model. However, these papers consider a different situation to the one being considered here. We want to estimate the marginal survival function of the failure time variable when there is informative censoring, whereas these papers estimate and compare the rate of change of a continuous variable measuring physiological function or disease status, when patients that discontinue from the study are considered to be informatively censored.

Koziol-Green estimators There is also a class of models in Braekers and Veraverbeke (2001, 2003, 2005, 2008) known as Koziol-Green models, where a censoring variable is assumed to have a hazard function that is proportional to the hazard function of the failure time variable. This means that the relationship between the survival functions for the two variables is

$$S_C(t) = S_T(t)^\rho$$

for some $\rho > 0$. This assumption is used in Braekers and Veraverbeke (2001,2003) and they refer to the censoring variable as an informative censoring variable. However, they also assume that the censoring variable is independent of the failure time variable. Therefore, this censoring variable is not truly informative. The term "partially informative censoring" used in Braekers and Veraverbeke (2005) is preferred when referring to this type of censoring.

Braekers and Veraverbeke (2008) consider a Koziol-Green type model when there is also dependence between the failure time and censoring variables. This is the situation that is of interest here. They use a copula function (see Section 2.7) to specify the joint distribution function of T and C. The copula-graphic estimators from Zheng and Klein (1995), which were considered in Section 2.9, are extended to the fixed design regression case. This is useful as it allows the incorporation of covariates, but is still not applicable

to the situation being considered here, as in the fixed design regression case, observations only occur at the fixed design points $x_1, \ldots, x_n$. This means that covariates can only take the predefined values $x_1, x_2, \ldots, x_n$.

3.3 Siannis (2004) and Siannis et al. (2005) Sensitivity Analyses

The methods that will now be considered allow the sensitivity of parameter estimates to informative censoring to be assessed. Firstly, sensitivity analyses for parametric survival models will be considered. One such approach is given in Siannis (2004) and Siannis et al. (2005). They use the same assumption about the conditional distribution of C given T to obtain equations for sensitivity analyses using parametric marginal distributions for T and C. Siannis et al. (2005) give the method for the simplest case where there is only one type of censoring in the data set and use exponential marginal distributions in their example. Siannis (2004) gives an extended version of the sensitivity analysis that allows for non-informative censoring as well as one type of informative censoring. It is necessary to use this extended version when applying the sensitivity analysis to the Liver Registration as we have non-informative end-of-study censoring as well as potentially informative censoring. The method used to derive the sensitivity analysis equations in this section will be given in more detail in Chapter 4, where the sensitivity analysis is extended to incorporate piecewise parametric models. In this section, we will only cover enough of the derivation of the method in the simplest case where there are only scalar parameters to illustrate how the sensitivity analysis equations were obtained. Weibull marginal distributions are used when applying this method to the Liver Registration data set.

The marginal density functions of T and C are given by $f_T(t, \theta)$ and $f_C(c, \gamma)$, where θ is the parameter of interest and γ will be treated as a nuisance parameter. This means there will also be corresponding hazard and survival functions for both T and C. The score and information functions for the marginal density functions are also required, for $f_T(t, \theta)$ these are defined by

$$s_T(t, \theta) = \frac{\partial}{\partial \theta} \log f_T(t, \theta) \quad \text{and} \quad i_\theta = \text{Var}_T\{s_T(T, \theta)\}.$$

We can define $s_C(c, \gamma)$ and i_γ similarly. The C variable here relates only to the potentially informative censoring as no parametric form is assumed for the non-informative censoring. As both informative censoring and non-informative censoring could be observed, the indicator variable $Z_i = I(Y_i = C_i)$ is required as well as $\Delta_i = I(Y_i = T_i)$.

The assumption that is used in Siannis (2004) and Siannis et al. (2005) to make the joint distribution of T and C identifiable is that the conditional distribution of C given T

is

$$f_C(c|t,\gamma,\delta,\theta) = f_C(c,\gamma + \delta i_\gamma^{-1/2} B(t,\theta)), \tag{3.11}$$

that is it has the same distribution as the marginal distribution of C but with the parameter dependent on $T = t$. The dependence is determined by δ and $B(t,\theta)$, where δ is a correlation coefficient and $B(t,\theta)$ is a bias function. The conditional density function in (3.11) can be approximated by

$$f_C(c|t,\gamma,\delta,\theta) \simeq f_C(c,\gamma)\left[1 + \delta i_\gamma^{-1/2} s_C(c,\gamma) B(t,\theta)\right]. \tag{3.12}$$

Let $\ell_\delta(\theta,\gamma)$, be the log-likelihood function when T and C are dependent as outlined above in Section 4.1. Then

$$\begin{aligned}\ell_\delta(\theta,\gamma) = \sum_{i=1}^{n} \Big\{ &\Delta_i \log K_1(t_i) + Z_i(1-\Delta_i)\log K_2(t_i) \\ &+ (1-\Delta_i)(1-Z_i)\log K_3(t_i)\Big\},\end{aligned} \tag{3.13}$$

where

$$\begin{aligned} K_1(t_i) &= \int_{t_i}^{\infty} f_{T,C}(t_i,u)du \\ K_2(t_i) &= \int_{t_i}^{\infty} f_{T,C}(u,t_i)du \\ \text{and} \quad K_3(t_i) &= \int_{t_i}^{\infty}\int_{t_i}^{\infty} f_{T,C}(t,c)dt\,dc. \end{aligned} \tag{3.14}$$

These can be thought of as the likelihood contributions for each of the three types of observations that may occur in each interval. The joint density function $f_{T,C}(t,c)$ is given by $f_T(t) f_C(c|t,\gamma,\delta,\theta)$ using the approximation of $f_C(c|t,\gamma,\delta,\theta)$ given in (3.12). When the forms of the contributions in (3.14) using this form of the joint density function are substituted in (3.13), then the log-likelihood becomes

$$\begin{aligned}\ell_\delta(\theta,\gamma) \simeq \ell_0(\theta,\gamma) - \delta i_\gamma^{-1/2} \sum_{i=1}^{n} \Big\{ &\Delta_i B(t_i,\theta)\frac{\partial}{\partial\gamma} H_C(t_i,\gamma) \\ &+ (1-\Delta_i)(1-Z_i)\frac{\partial}{\partial\gamma} H_C(t_i,\gamma)\mu(t_i,\theta) \\ &- Z_i(1-\Delta_i)s_C(t_i,\gamma)\mu(t_i,\theta)\Big\},\end{aligned} \tag{3.15}$$

where

$$\mu(t_i,\theta) = \frac{\displaystyle\int_{t_i}^{\infty} B(u,\theta) f_T(u,\theta)du}{S_T(t_i,\theta)}.$$

For a fixed value of δ, $\hat{\theta}_\delta$ is the value that maximises (3.15). The first term in (3.15), $\ell_0(\theta,\gamma)$, is the log-likelihood under the assumption that T and C are independent. This log likelihood is used to find the maximum likelihood estimates (MLEs), $\hat{\theta}_0$ and $\hat{\gamma}_0$.

The aim of this method is to approximate the value of $\hat{\theta}_\delta - \hat{\theta}_0$, which is done by rearranging Taylor expansions of the score functions

$$r_0(\hat{\theta}_0) = \frac{\partial}{\partial\theta}\ell_0(\theta,\gamma)\Big|_{\hat{\theta}_0} \quad \text{and} \quad r_\delta(\hat{\theta}_\delta) = \frac{\partial}{\partial\theta}\ell_\delta(\theta,\gamma)\Big|_{\hat{\theta}_\delta}. \tag{3.16}$$

The score functions given in (3.16) are expanded about θ and set equal to zero to give

$$\begin{aligned} r_0(\hat{\theta}_0) &\simeq r_0(\theta) - (\hat{\theta}_0 - \theta)i(\theta) = 0 \\ r_\delta(\hat{\theta}_\delta) &\simeq r_\delta(\theta) - (\hat{\theta}_\delta - \theta)i(\theta) = 0 \end{aligned} \tag{3.17}$$

where

$$i(\theta) = -\frac{\partial^2}{\partial\theta^2}\ell_0(\theta,\gamma).$$

Rearranging the two equations in (3.17) gives

$$(\hat{\theta}_\delta - \hat{\theta}_0)i(\theta) \simeq r_\delta(\theta) - r_0(\theta).$$

So, an approximation of the difference between the parameter estimates is given by

$$\begin{aligned} \hat{\theta}_\delta - \hat{\theta}_0 \simeq \delta i_\gamma^{-1/2}(i(\theta))^{-1} \sum_{i=1}^{n} \Big\{ & Z_i(1-\Delta_i)s_C(t_i,\gamma)\frac{\partial\mu(t_i,\theta)}{\partial\theta} \\ & - (1-Z_i)(1-\Delta_i)\frac{\partial H_C(t_i,\gamma)}{\partial\gamma}\frac{\partial\mu(t_i,\theta)}{\partial\theta} \\ & - \Delta_i\frac{\partial H_C(t_i,\gamma)}{\partial\gamma}\frac{\partial B(t_i,\theta)}{\partial\theta}\Big\} \end{aligned} \tag{3.18}$$

We can see that in (3.18) there are parameter estimates on the LHS of the approximation and parameters on the RHS. This is a consequence of rearranging the Taylor expansions of the score functions, which are given in (3.17). So when the sensitivity analysis is applied, the parameters on the RHS of (3.18) must be replaced by estimated values.

Before the sensitivity analysis can be applied a form of the bias function $B(t,\theta)$ needs to be chosen. A detailed explanation of the bias function chosen is given in Section 4.2.1. The bias function we use is the same as the one used in Siannis et al. (2005). The expression in (3.18) can also be simplified by assuming a proportional hazards structure. This is discussed in Section 4.2.2 and the same structure is used in Siannis et al. (2005). After these changes, the expression in (3.18) becomes

$$\hat{\theta}_\delta - \hat{\theta}_0 \simeq \delta i(\theta)^{-1} \sum_{i=1}^{n} \big\{ H_T(t_i,\theta)H_C(t_i,\gamma) - Z_i(1-\Delta_i)H_T(t_i,\theta)\big\}, \tag{3.19}$$

where

$$i(\theta) = \sum_{i=1}^{n} H_T(t_i,\theta).$$

Siannis et al. (2005) include covariates in the sensitivity analysis by replacing θ and γ by the linear predictors $w(\mathbf{x}) = \boldsymbol{\theta}'\mathbf{x}$ and $z(\mathbf{x}) = \boldsymbol{\gamma}'\mathbf{x}$. They derive an expression for the sensitivity analysis that approximates the difference between the vectors of parameter estimates $\hat{\boldsymbol{\theta}}_\delta$ and $\hat{\boldsymbol{\theta}}_0$. The vector $\hat{\boldsymbol{\theta}}_\delta$ maximises the log-likelihood

$$\begin{aligned}\ell_\delta(\boldsymbol{\theta},\boldsymbol{\gamma}) \simeq \ell_0(\boldsymbol{\theta},\boldsymbol{\gamma}) - \delta i_{\boldsymbol{\gamma}}^{-1/2} \sum_{i=1}^{n} \Big\{ & \Delta_i B(t_i,\boldsymbol{\theta},\mathbf{x}) \frac{\partial}{\partial \boldsymbol{\gamma}} H_C(t_i,\boldsymbol{\gamma},\mathbf{x}) \\ & + (1-\Delta_i)(1-Z_i) \frac{\partial}{\partial \boldsymbol{\gamma}} H_C(t_i,\boldsymbol{\gamma},\mathbf{x}) \mu(t_i,\boldsymbol{\theta},\mathbf{x}) \\ & - Z_i(1-\Delta_i) s_C(t_i,\boldsymbol{\gamma},\mathbf{x}) \mu(t_i,\boldsymbol{\theta},\mathbf{x}) \Big\}, \end{aligned} \tag{3.20}$$

where

$$\mu(t_i,\boldsymbol{\theta},\mathbf{x}) = \frac{\displaystyle\int_{t_i}^{\infty} B(u,\boldsymbol{\theta},\mathbf{x}) f_T(u,\boldsymbol{\theta},\mathbf{x}) du}{S_T(t_i,\boldsymbol{\theta},\mathbf{x})}.$$

The log-likelihood in (3.20) is the log-likelihood in (3.15) that has been extended to allow the inclusion of covariates and vectors of parameters. Similarly $\hat{\boldsymbol{\theta}}_0$ is the vector of parameter estimates that maximises $\ell_0(\boldsymbol{\theta},\boldsymbol{\gamma})$. We shall express $\hat{\boldsymbol{\theta}}_\delta - \hat{\boldsymbol{\theta}}_0$ using slightly different notation to Siannis et al. (2005). The equation

$$\hat{\boldsymbol{\theta}}_\delta - \hat{\boldsymbol{\theta}}_0 \simeq \delta i(\boldsymbol{\theta},\mathbf{x})^{-1} (\mathbf{r}_\delta(\boldsymbol{\theta}) - \mathbf{r}_0(\boldsymbol{\theta})), \tag{3.21}$$

is found by rearranging vectorised versions of the Taylor expansions in (3.17). The kth component of $\mathbf{r}_\delta(\boldsymbol{\theta}) - \mathbf{r}_0(\boldsymbol{\theta})$ is

$$\sum_{i=1}^{n} x_{ik} \big\{ H_T(t_i,\boldsymbol{\theta},\mathbf{x}_i) H_C(t_i,\boldsymbol{\gamma},\mathbf{x}_i) - Z_i(1-\Delta_i) H_T(t_i,\boldsymbol{\theta},\mathbf{x}_i) \big\}. \tag{3.22}$$

The information matrix is now $i(\boldsymbol{\theta},\mathbf{x})$, where the (k,l)th element is given by

$$-\frac{\partial}{\partial \theta_k} \frac{\partial}{\partial \theta_l} \ell_0(\boldsymbol{\theta},\boldsymbol{\gamma},\mathbf{x}) = x_{ik} x_{il} H_T(t_i,\boldsymbol{\theta},\mathbf{x}_i)$$

However, Siannis et al. (2005) did not use (3.21) when applying the sensitivity analysis to data. Instead they performed the sensitivity analysis on $w(\mathbf{x})$ rather than $\boldsymbol{\theta}$. However, when applying the sensitivity analysis to the Liver Registration data set we shall the perform the sensitivity analysis for $\boldsymbol{\theta}$ as well as that for $w(\mathbf{x})$.

The sensitivity analysis equation for performing the sensitivity analysis equation on $w(\mathbf{x})$ is now derived. The quantity of interest is now $\hat{w}_\delta(\mathbf{x}) - \hat{w}_0(\mathbf{x})$ where $\hat{w}_\delta(\mathbf{x})$ is the estimated linear predictor using the vector $\hat{\boldsymbol{\theta}}_\delta$ that maximises the log-likelihood in (3.20). Similarly $\hat{w}_0(\mathbf{x})$ is the estimated linear predictor using the vector $\hat{\boldsymbol{\theta}}_0$ that maximises $\ell_0(\boldsymbol{\theta},\boldsymbol{\gamma})$. The linear predictors $w(\mathbf{x})$ and $z(\mathbf{x})$ are treated as scalar quantities so the

sensitivity analysis can be found by replacing θ and γ in (3.18) by $w(\mathbf{x})$ and $z(\mathbf{x})$. The sensitivity analysis equation becomes

$$\hat{w}_\delta(\mathbf{x}) - \hat{w}_0(\mathbf{x}) \simeq \delta i(w(\mathbf{x}))^{-1} \sum_{i=1}^{n} \big\{ H_T(t_i, w(\mathbf{x})) H_C(t_i, z(\mathbf{x})) \\ - Z_i(1-\Delta_i) H_T(t_i, w(\mathbf{x})) \big\}, \tag{3.23}$$

where

$$i(w(\mathbf{x})) = -\frac{\partial}{\partial w(\mathbf{x})} \ell_0(w(\mathbf{x}), z(\mathbf{x})) = \sum_{i=1}^{n} H_T(t_i, w(\mathbf{x})).$$

We can see that (3.23) only applies to the covariate vector $\mathbf{x}$, so to estimate the change in linear predictors for all individuals in the data set, all observed covariate vectors must be considered. This dependence on the covariate vector $\mathbf{x}$ also means that the same covariates have to be included in both the model for time to death and the model for time to censoring.

We can see that in (3.21) and (3.23) there is the same issue that was observed in (3.18). There are parameter estimates on the LHS of the expressions and parameters on the RHS. This means that the parameters need to be replaced by estimated values when applying the sensitivity analysis.

3.3.1 Comparison with Scharfstein and Robins (2002)

In this section the assumption in (3.11) will be compared to the assumption used in Scharfstein and Robins (2002). The aim of this is to make the interpretation of the assumption in (3.11) easier to understand. Scharfstein and Robins (2002) assume that the censoring process follows a proportional hazards model, so that the conditional hazard function for C can be expressed as

$$h_C(c|T, T > c) = h_{C0}(c) \exp(q(c, T)), \tag{3.24}$$

that is the conditional hazard for C given T is the baseline hazard multiplied by a function of T. The function $q(c, T)$ quantifies the dependence between T and C just after time c, for those who are still at risk at time c. This "censoring bias function" determines the way T enters the proportional hazards model for the cause-specific hazard of censoring.

So that the two assumptions can be compared, the corresponding conditional hazard function for the conditional density function in (3.11) needs to be found. The form of this conditional hazard function is given in Siannis et al (2005) and we shall now give the derivation of this term. Firstly, we use that

$$S_C(c|T, \gamma, \delta, \theta) = \int_c^\infty f_C(c|T, \gamma, \delta, \theta) dc \\ \simeq S_C(c, \gamma)[1 - \delta i_\gamma^{-1/2} B(t, \theta) \frac{\partial}{\partial \gamma} H_C(c, \gamma)],$$

which means that the conditional hazard can be expressed as

$$\begin{aligned} h_C(c|T,\gamma,\delta,\theta) &= -\frac{\partial}{\partial c}\log S_C(c|T,\gamma,\delta,\theta) \\ &\simeq -\frac{\partial}{\partial c}\left[\log S_C(c,\gamma) + \log\left(1-\delta i_\gamma^{-1/2}B(t,\theta)\frac{\partial}{\partial\gamma}H_C(c,\gamma)\right)\right] \end{aligned} \tag{3.25}$$

The approximation $\log(1+x) \simeq x$ is used to simplify the second term in (3.25), so that the conditional hazard becomes

$$h_C(c|T,\gamma,\delta,\theta) \simeq -\frac{\partial}{\partial c}\log S_C(c,\gamma) - \frac{\partial}{\partial c}\left(-\delta i_\gamma^{-1/2}B(t,\theta)\frac{\partial}{\partial\gamma}H_C(c,\gamma)\right).$$

This can be rearranged to give

$$h_C(c|T,\gamma,\delta,\theta) \simeq h_C(c,\gamma)\left[1+\delta i_\gamma^{-1/2}B(t,\theta)\frac{\partial}{\partial\gamma}\log h_C(c,\gamma)\right]. \tag{3.26}$$

To be able to compare (3.11) with (3.24), the conditional hazard in (3.11) needs to be expressed as a proportional hazards model, with the baseline hazard function being multiplied by some function. To do this, the approximation $e^x \simeq 1+x$ is used in (3.26), so that the conditional hazard function is now

$$h_C(c,\gamma)\exp\left(\delta i_\gamma^{-1/2}B(T,\theta)\frac{\partial}{\partial\gamma}\log h_C(c,\gamma)\right). \tag{3.27}$$

If (3.27) is compared with (3.24), then we can see the two hazard functions have a similar form. The baseline hazard in (3.24), has been replaced with a parametric baseline hazard in (3.27). Also, we see that the specification of $q(c,T)$ in (3.24) is the same as choosing $\delta B(T,\theta)$ in (3.27). This means that $\delta B(T,\theta)$ also quantifies the dependence between T and C just after time c and determines the way that T enters the proportional hazards model for censoring.

3.3.2 Application to the Liver Registration data set

This sensitivity analysis is now applied to the Liver Registration data set. Firstly, the sensitivity analysis will be performed on $w(\mathbf{x})$ and then the sensitivity analysis for $\boldsymbol{\theta}$ will be applied. Siannis et al. (2005) assumed exponential marginal models for T and C and Siannis (2004) used Weibull marginal models for T and C. When applying this method to the Liver Registration data set, Weibull marginal models are used as these are more flexible than exponential marginal models.

When applying the sensitivity analysis to $w(\mathbf{x})$, the marginal density functions are given by

$$\begin{aligned} f_T(t,w(\mathbf{x}),\eta_T) &= e^{w(\mathbf{x})}\eta_T t^{\eta_T-1}\exp(-e^{w(\mathbf{x})}t^{\eta_T}) \quad \text{and} \\ f_C(t,z(\mathbf{x}),\eta_C) &= e^{z(\mathbf{x})}\eta_C t^{\eta_C-1}\exp(-e^{z(\mathbf{x})}t^{\eta_C}). \end{aligned}$$

This means that the integrated hazard functions are

$$H_T(t, w(\mathbf{x}), \eta_T) = e^{w(\mathbf{x})} t^{\eta_T} \quad \text{and}$$
$$H_C(t, z(\mathbf{x}), \eta_C) = e^{z(\mathbf{x})} t^{\eta_C}. \tag{3.28}$$

Here the scale parameters $w(\mathbf{x})$ and $z(\mathbf{x})$ are linear predictors that incorporate the following covariates: age at registration, ethnicity, primary liver disease category and UKELD score at registration. The same covariates need to be included in the models for time to death and time to censoring for this sensitivity analysis.

The sensitivity analysis will be conducted on the scale parameter for T, $w(\mathbf{x})$, as this is the parameter of interest and the shape parameters, η_T and η_C, are treated as nuisance parameters. The scale parameter for C, $z(\mathbf{x})$, is also treated as a nuisance parameter. If the integrated hazards in (3.28) are substituted into (3.23) then the sensitivity analysis equation becomes

$$\hat{w}_\delta(\mathbf{x}) - \hat{w}_0(\mathbf{x}) \simeq \delta \frac{\sum_{i=1}^n \left\{ e^{z(\mathbf{x})} t_i^{\eta_T + \eta_C} - Z_i(1 - \Delta_i) t_i^{\eta_T} \right\}}{\sum_{i=1}^n t_i^{\eta_T}}. \tag{3.29}$$

This can be thought of as δ multiplied by a sensitivity index, U. As in (3.23), we have parameter estimates on the LHS of (3.29) and parameters on the RHS. To overcome this issue when applying the sensitivity analysis, $z(\mathbf{x})$, η_T and η_C are replaced by their estimates from the Weibull proportional hazards model that assumes non-informative censoring. It is found that $\hat{\eta}_{T0} = 1.03$ and $\hat{\eta}_{C0} = 0.9297$. The estimate for η_T was not found to be significantly different from one so an exponential model could be used for T, however the estimate for η_C was significantly different from one so the use of Weibull marginal models is justified.

As there are many different combinations of the covariates in the Liver Registration data set, $\hat{z}_0(\mathbf{x})$ takes a range of values so the sensitivity index needs to be computed over this range. The easiest way of displaying the results is to plot δU over the range of $\hat{z}_0(\mathbf{x})$, which is shown in Figure 3.9 for $\delta = 0.2$ and 0.3. The range of values for $\hat{z}_0(\mathbf{x})$ used on the horizontal axis in Figure 3.9 is the observed range of $\hat{z}_0(\mathbf{x})$ for the Liver Registration data set. The largest values of $\hat{w}_\delta(\mathbf{x}) - \hat{w}_0(\mathbf{x})$ are observed for patients with the largest values of $\hat{z}_0(\mathbf{x})$. These are the patients which have the greatest hazard of censoring. We see that for these individuals, the change in the estimated linear predictor seems large enough that results obtained assuming non-informative censoring could be misleading. However, to be sure of this the effect on a value of interest, such as the survival function of individuals in the data set, should be examined. When we apply the sensitivity analysis derived in Chapter 4 to the Liver Registration data set, this will be investigated.

The sensitivity analysis for $\boldsymbol{\theta}$ will now be applied to the Liver Registration data set. Again Weibull marginal models are assumed for T and C. For simplicity, $z(\mathbf{x})$ will be used

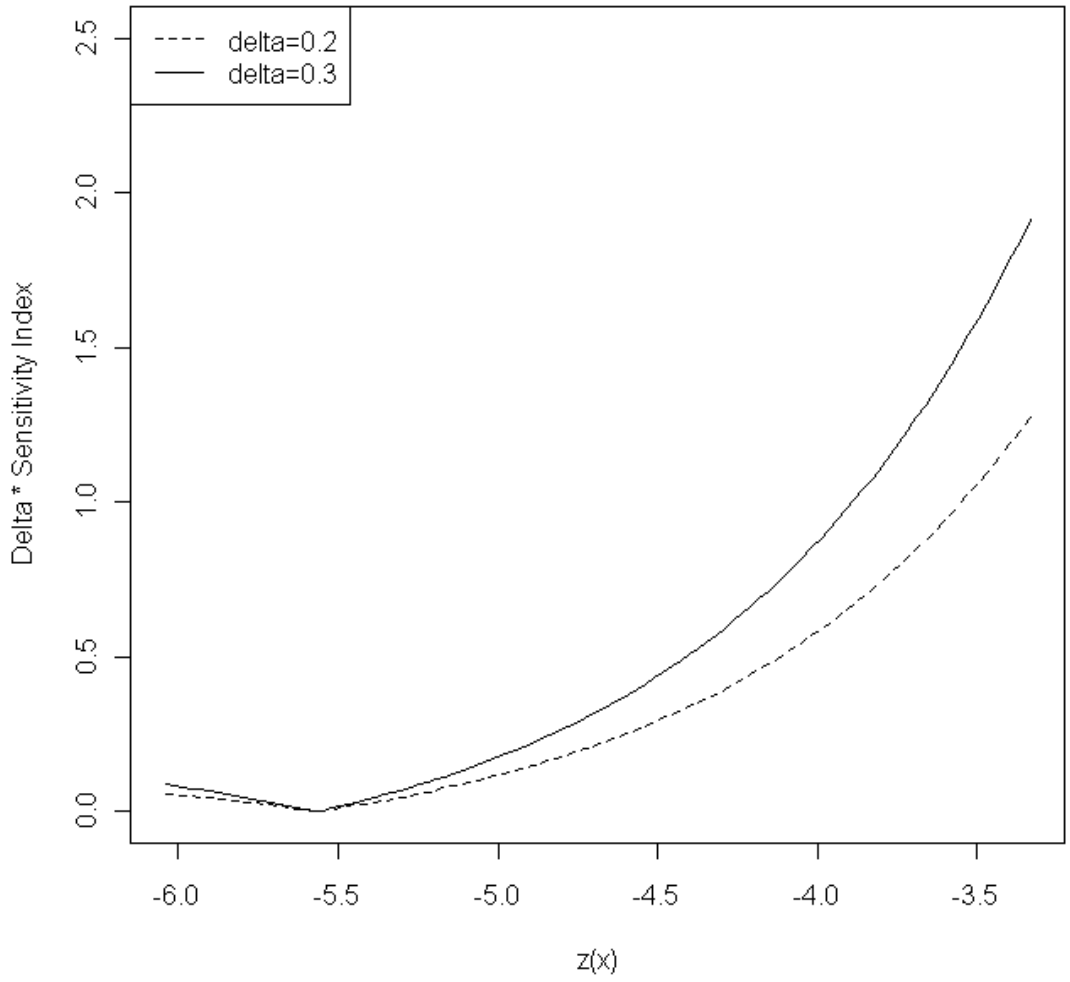

Figure 3.9: Plot showing δ times the sensitivity index, U, over the range of observed values for $\hat{z}_0(\mathbf{x})$ for the individuals in the Liver registration data set, using $\delta = 0.2$ and 0.3.

as the scale parameter for C, rather than the vector $\boldsymbol{\gamma}$. This means the marginal density functions are now given by

$$f_T(t, \boldsymbol{\theta}, \mathbf{x}, \eta_T) = e^{\boldsymbol{\theta}'\mathbf{x}} \eta_T t^{\eta_T - 1} \exp(-e^{\boldsymbol{\theta}'\mathbf{x}} t^{\eta_T}) \quad \text{and}$$
$$f_C(t, z(\mathbf{x}), \eta_C) = e^{z(\mathbf{x})} \eta_C t^{\eta_C - 1} \exp(-e^{z(\mathbf{x})} t^{\eta_C}).$$

The integrated hazard functions are now

$$H_T(t, \boldsymbol{\theta}, \mathbf{x}, \eta_T) = e^{\boldsymbol{\theta}'\mathbf{x}} t^{\eta_T} \quad \text{and}$$
$$H_C(t, z(\mathbf{x}), \eta_C) = e^{z(\mathbf{x})} t^{\eta_C}. \tag{3.30}$$

It is the vector of parameters for T, $\boldsymbol{\theta}$, that is of interest. So, η_T, η_C and $z(\mathbf{x})$ will again be treated as nuisance parameters. For notational simplicity, it is assumed that the same covariate vector is used in both the model for time to death and the model for time to censoring. However, it is not a requirement for this sensitivity analysis. Therefore, age, ethnicity, primary liver disease category and UKELD score are used in the model for time to death and primary liver disease category, UKELD score, height and blood group are used in the model for time to censoring.

The sensitivity analysis equation in (3.21) will be used to carry out the sensitivity analysis for $\boldsymbol{\theta}$. When substituting the integrated hazard functions in (3.30) into (3.22), the expression for the kth component of $\mathbf{r}_\delta(\boldsymbol{\theta}) - \mathbf{r}_0(\boldsymbol{\theta})$ becomes

$$\sum_{i=1}^{n} x_{ik} \left\{ e^{\boldsymbol{\theta}'\mathbf{x}_i} e^{z(\mathbf{x}_i)} t_i^{\eta_T + \eta_C} - Z_i(1 - \Delta_i) e^{\boldsymbol{\theta}'\mathbf{x}_i} t_i^{\eta_T} \right\}. \tag{3.31}$$

and the (k, l)th element of the information matrix $i(\boldsymbol{\theta}, \mathbf{x})$ in (3.21) becomes

$$\sum_{i=1}^{n} x_{ik} x_{il} e^{\boldsymbol{\theta}'\mathbf{x}_i} t_i^{\eta_T}.$$

We can see that in (3.31) we have the parameter vector $\boldsymbol{\theta}$ as well as $z(\mathbf{x})$, η_T and η_C. These all need to be replaced with their estimates from the Weibull proportional hazards model that assumes non-informative censoring.

Table 3.1 shows the estimated values of the components of $\hat{\boldsymbol{\theta}}_\delta - \hat{\boldsymbol{\theta}}_0$ for $\delta = 0.2$ and $\delta = 0.3$. We see that for some covariates there are positive changes in the parameter estimates, while others have negative changes in the parameter estimates. Positive values in Table 3.1 mean that the element of $\hat{\boldsymbol{\theta}}_\delta$ for that covariate is larger than the corresponding element of $\hat{\boldsymbol{\theta}}_0$. So, this suggests that the hazard ratio of the covariate is being underestimated by the model assuming non-informative censoring. Conversely, negative values in Table 3.1 mean that the parameter estimate for the covariate from the model assuming informative censoring is smaller than the corresponding parameter estimate from the model assuming

non-informative censoring. Therefore, the sensitivity analysis is suggesting that the hazard ratio for these covariates are overestimated by the model that assumes $\delta = 0$.

So, the sensitivity analysis for $\boldsymbol{\theta}$ suggests the hazard ratio for patients with hepatitis B virus infection is being underestimated, whereas the hazard ratios for patients with other levels of primary liver disease are being overestimated. Patients of either white or black ethnic origin are having their hazard ratios overestimated, whereas the hazard ratios for patients of asian or oriental ethnic origin are being underestimated. The sensitivity analysis also suggests that the hazard ratios for UKELD score and age are being slightly overestimated by the model that assumes non-informative censoring.

The effects that the estimated changes in Table 3.1 have on the parameter estimates are shown in Table 3.2. The p-values of the estimates are also shown. These are all calculated using the standard errors of the estimates from the model assuming non-informative censoring. This can be done as Siannis et al. (2005) show that

$$\{\mathrm{Var}(\hat{\theta}_\delta)\}^{1/2} \simeq \{\mathrm{Var}(\hat{\theta}_0)\}^{1/2} + O(\delta^2).$$

Only linear values of δ are considered in the sensitivity analysis so the standard error of the parameter estimate from the model assuming informative censoring can be approximated by the standard error of the parameter estimate from the model assuming non-informative censoring. This approximation should only be used if the value of δ is fairly small.

3.4 Zhang and Heitjan (2006) Sensitivity Analysis

An alternative sensitivity analysis for parametric survival models is presented in Zhang and Heitjan (2006). Again, the marginal density functions of T and C are given by $f_T(t, \theta)$ and $f_C(c, \gamma)$, where θ is the parameter of interest and γ will be treated as a nuisance parameter. Non-informative censoring could also be observed but no parametric distribution will be assumed for this type of censoring. Therefore, for simplicity we use C to denote the informative censoring. As there are several types of censoring that can be observed in addition to the failure time then two indicator variables are required to distinguish between the events. These are $\Delta_i = I(Y_i = T_i)$ and $Z_i = I(Y_i = C_i)$. The likelihood function that incorporates one type of informative censoring as well as non-informative censoring is

$$L_\delta(\theta, \gamma) = \prod_{i=1}^{n} \mathrm{Int}_{1i}^{\Delta_i(1-Z_i)} \mathrm{Int}_{2i}^{Z_i(1-\Delta_i)} \mathrm{Int}_{3i}^{(1-\Delta_i)(1-Z_i)} \tag{3.32}$$

Parameter	$\hat{\boldsymbol{\theta}}_{0.2} - \hat{\boldsymbol{\theta}}_0$	$\hat{\boldsymbol{\theta}}_{0.3} - \hat{\boldsymbol{\theta}}_0$
Intercept	0.5529	0.8293
PLD - PBC	-0.0717	-0.1076
PLD - PSC	-0.0815	-0.1223
PLD - ALD	-0.0343	-0.0515
PLD - AID	-0.0675	-0.1013
PLD - HCV	-0.0694	-0.1040
PLD - HBV	0.1122	0.1683
PLD - Cancer	-0.0675	-0.1012
PLD - Metabolic	-0.00001	-0.00002
PLD - Other	-0.1027	-0.1540
Ethnicity - White	-0.0194	-0.0292
Ethnicity - Asian	0.0325	0.0487
Ethnicity - Black	-0.0379	-0.0569
Ethnicity - Chinese	0.0721	0.1082
UKELD score	-0.0046	-0.0069
Age	-0.0012	-0.0018

Table 3.1: The results of the Siannis sensitivity analysis for $\boldsymbol{\theta}$ using Weibull marginals. The table shows each component of the vector $\hat{\boldsymbol{\theta}}_\delta - \hat{\boldsymbol{\theta}}_0$ for $\delta = 0.2$ and $\delta = 0.3$.

Parameter	$\hat{\boldsymbol{\theta}}_0$	p-value	$\hat{\boldsymbol{\theta}}_{0.2}$	p-value	$\hat{\boldsymbol{\theta}}_{0.3}$	p-value
Intercept	-20.6134	< 0.001	-20.0605	< 0.001	-19.7841	< 0.001
PLD - PBC	-0.2260	0.508	-0.2977	0.383	-0.3336	0.328
PLD - PSC	-0.9060	0.022	-0.9875	0.013	-1.0283	0.010
PLD - ALD	-0.4644	0.138	-0.4987	0.111	-0.5159	0.099
PLD - AID	-0.0141	0.966	-0.0817	0.806	-0.1154	0.729
PLD - HCV	0.2715	0.409	0.2022	0.538	0.1675	0.610
PLD - HBV	-0.4724	0.418	-0.3602	0.537	-0.3041	0.602
PLD - Cancer	-1.4244	0.062	-1.5019	0.050	-1.5357	0.046
PLD - Metabolic	0.6656	0.063	0.6656	0.063	0.6656	0.063
PLD - Other	0.3657	0.282	0.2630	0.439	0.2117	0.534
Ethnicity - White	0.9596	0.342	0.9401	0.351	0.9304	0.356
Ethnicity - Asian	-0.0369	0.972	-0.0044	0.997	0.0118	0.991
Ethnicity - Black	0.9273	0.408	0.8894	0.428	0.8704	0.438
Ethnicity - Chinese	-0.7135	0.619	-0.6413	0.655	-0.6053	0.673
UKELD score	0.1943	< 0.001	0.1898	< 0.001	0.1875	< 0.001
Age	0.0308	< 0.001	0.0296	< 0.001	0.0290	< 0.001

Table 3.2: The approximate values of the vectors of parameter estimates $\hat{\boldsymbol{\theta}}_{0.2}$ and $\hat{\boldsymbol{\theta}}_{0.3}$ obtained using the results of the Siannis sensitivity analysis given in Table 3.1. The parameter estimates of the model assuming non-informative censoring are included for comparison. The p-values of the parameter estimates are also included, they were calculated using the standard errors of the model assuming non-informative censoring.

where

$$\text{Int}_{1i} = \int_{t_i}^{\infty} f_T(t_i, \theta) f_{C|T}(u|t_i, \gamma, \delta) du$$
$$\text{Int}_{2i} = \int_{t_i}^{\infty} f_T(u, \theta) f_{C|T}(t_i|u, \gamma, \delta) du \quad \text{and}$$
$$\text{Int}_{3i} = \int_{t_i}^{\infty} \int_{t_i}^{\infty} f_T(u, \theta) f_{C|T}(v|u, \gamma, \delta) du dv.$$

So the conditional distribution of C given T needs to be specified so that this likelihood is well defined. As in Section 3.3 this conditional distribution is assumed to be the same as the distribution of C but with the parameter allowed to depend on t. However in Zhang and Heitjan (2006), δt replaces $\delta i_\gamma^{-1/2} B(t, \theta)$ in (3.11).

To evaluate the sensitivity of an estimate of θ to small departures of δ from zero, the rate at which $\hat{\theta}_\delta$ departs from $\hat{\theta}_0$ as δ varies from zero needs to be calculated. Troxel et al. (2004) gave an index of sensitivity to non-ignorability (ISNI), which is given by

$$\text{ISNI}(\hat{\theta}) = \left.\frac{\partial \hat{\theta}_\delta}{\partial \delta}\right|_{\delta=0} = -\left[\frac{\partial^2 \ell_\delta}{\partial \theta \partial \theta'}\right]^{-1} \left.\frac{\partial^2 \ell_\delta}{\partial \theta \partial \delta}\right|_{\delta=0, \hat{\gamma}_0, \hat{\theta}_0} \tag{3.33}$$

where ℓ_δ is the logarithm of the likelihood in (3.32).

As ISNI is the derivative of $\hat{\theta}$ with respect to δ, the value of $\hat{\theta}$ for a fixed value of δ is approximately

$$\hat{\theta}_\delta \simeq \hat{\theta}_0 + \delta \text{ISNI}(\hat{\theta}). \tag{3.34}$$

When applying the sensitivity analysis described in Section 3.3, the value of $\hat{\theta}$ for a given value of δ can be approximated by

$$\hat{\theta}_\delta \simeq \hat{\theta}_0 + \delta U, \tag{3.35}$$

so it is possible to compare the values of $\text{ISNI}(\hat{\theta})$ and U.

Zhang and Heitjan (2006) suggest a method to assess whether $\hat{\theta}$ is sensitive to the informative censoring. They define the inference to be affected by informative censoring if the estimate changes by more than 1 standard error (SE) of the parameter estimate under the model that assumes non-informative censoring. The value of δ that causes a change of 1 SE in $\hat{\theta}$ is

$$\delta^* = \frac{\text{SE}(\hat{\theta}_0)}{\text{ISNI}(\hat{\theta})}$$

and this is considered to be the smallest value of δ that causes a substantial change in $\hat{\theta}$. They say that the plausibility of δ^* can be checked by plotting t^* against a suitable measure, where t^* varies over the range of observed values. A suitable measure is one for which the plausibility is easily assessed, such as the mean or hazard function of C given T. If the value of δ^* is deemed to be plausible, then $\hat{\theta}$ is sensitive to the informative censoring

in the data set. However, no criteria are given to establish whether the values of the measure are indeed plausible or not. For example, in one application considered in Zhang and Heitjan (2006) they say that the value of δ^* is plausible because the hazard ratios do not vary by more than two. However, the choice of this value seems rather arbitrary so this method does not seem to be that useful when establishing whether $\hat{\theta}$ is sensitive to informative censoring.

3.4.1 Application to the Liver Registration data set

Only exponential marginal models for T and C are considered in Zhang and Heitjan (2006) and the form of ISNI is given just for the situation where there is only informative censoring considered. As the Liver Registration data set also has some non-informative censoring, the form of ISNI given in Zhang (2004) is used, which is

$$\text{ISNI}(\hat{\theta}) = \left\{ e^{\hat{\theta}_0} \sum_{i=1}^{n} t_i \right\}^{-1} \times e^{-\hat{\theta}_0} \left[\sum_{i=1}^{n} (1-\Delta_i)(1-Z_i) e^{\hat{\gamma}_0} t_i - Z_i(1-\Delta_i)(1-t_i e^{\hat{\gamma}_0}) \right] \tag{3.36}$$

where

$$e^{\hat{\theta}_0} = \frac{\sum_{i=1}^{n} \Delta_i (1-Z_i)}{\sum_{i=1}^{n} t_i} \qquad e^{\hat{\gamma}_0} = \frac{\sum_{i=1}^{n} Z_i (1-\Delta_i)}{\sum_{i=1}^{n} t_i}.$$

If (3.36) is used to calculate the ISNI for the Liver Registration data set when assuming exponential marginal distributions with scalar parameters, then $\text{ISNI}(\hat{\theta}) = -757.72$. Zhang and Heitjan (2006) observed values of $\text{ISNI}(\hat{\theta})$ of a similar magnitude when they applied the sensitivity analysis to their data. This value of $\text{ISNI}(\hat{\theta})$ can be compared to the value of U obtained when applying the Siannis sensitivity analysis to the data set with the same marginal distributions assumed, where

$$U = \frac{\sum_{i=1}^{n} e^{\hat{\gamma}_0} t_i^2 - Z_i(1-\Delta_i t_i)}{\sum_{i=1}^{n} t_i}.$$

This gives $U = 1.05$, which gives very different results to ISNI. If a positive value of δ is assumed, which assumes positive dependence between T and C, then the value of ISNI suggests there will be a very large decrease in the parameter estimate, whereas U suggests there will be a moderate increase in the parameter estimate. The result from the Siannis sensitivity analysis seems more realistic, as if we assume positive dependence between T and C and those being censored have a lower expected survival, then a model that incorporates this dependence should give a value of $E(T) = e^{-\theta}$ that is lower than under the model that assumes non-informative censoring. This means an increase in the parameter estimate under the assumption of informative censoring is expected. Observing

such unexpected values of ISNI($\hat{\theta}$) suggests that there could be an error in the method presented in Zhang and Heitjan (2006). However, we were unable to find any errors in the derivation of (3.36).

The values of δ^* that would give a change in the parameter estimate of one standard error of $\hat{\theta}$ under the non-informative censoring model can also be computed and compared. For the Zhang and Heitjan sensitivity analysis, this is $\delta^* = -0.000028$, and for the Siannis sensitivity analysis, this is $\delta^* = 0.021$. While both these values are small, the δ^* for the Zhang and Heijtan sensitivity analysis seems unfeasibly small. However, use of this method to determine whether $\hat{\theta}$ is sensitive to informative censoring is not recommended as the method used to determine whether δ^* is plausible still seems to be subjective.

Covariates are not incorporated in (3.36), although if $\hat{\theta}_0$ and $\hat{\gamma}_0$ are replaced by $\hat{w}_0(\mathbf{x}) = \hat{\boldsymbol{\theta}}_0'\mathbf{x}$ and $\hat{z}_0(\mathbf{x}) = \hat{\boldsymbol{\gamma}}_0'\mathbf{x}$ then the value of ISNI can be calculated over the ranges of $\hat{w}_0(\mathbf{x})$ and $\hat{z}_0(\mathbf{x})$. The covariates included in the vector $\mathbf{x}$ are the same as those used when applying the sensitivity analysis for $w(\mathbf{x})$ from Siannis et al. (2005) to the Liver Registration data set. These are recipient age, recipient ethnicity, primary liver disease category and UKELD score. The range of values that ISNI takes when including covariates is shown in the plot in Figure 3.10 for $\delta = 0.2$.

In Figure 3.10, we see that for some parameter combinations, the expected increase in the parameter estimate is observed, but again this tends to have an extremely large magnitude that does not seem feasible in reality. However, Figure 3.10 looks at all combinations of $\hat{w}_0(\mathbf{x})$ and $\hat{z}_0(\mathbf{x})$ for their observed ranges, when in the Liver Registration data set only some of these combinations are observed. Figure 3.11 shows the observed combinations of $\hat{w}_0(\mathbf{x})$ and $\hat{z}_0(\mathbf{x})$ for all the patients in the Liver Registration data set. We can see that none of the individuals have the combination of $\hat{w}_0(\mathbf{x})$ and $\hat{z}_0(\mathbf{x})$ that gives the largest value of ISNI. However, if the value of ISNI is calculated for each observed combination in the Liver Registration data set, it is found that it takes values in the interval (-24663436,4102221). The boundaries of this interval still have such large magnitudes that these values do not seem realistic.

3.5 Huang and Zhang (2008) Sensitivity Analysis

The previous sections 3.3 and 3.4 considered sensitivity analyses for parametric survival models. In this section and the following section 3.6, sensitivity analyses that use the Cox proportional hazards model for the marginal distributions are considered.

The model presented in this section extends the copula approach of Zheng and Klein (1994) to develop an estimation method for the bivariate proportional hazards model for competing risks. Marginally, each one of the dependent competing risks under study is

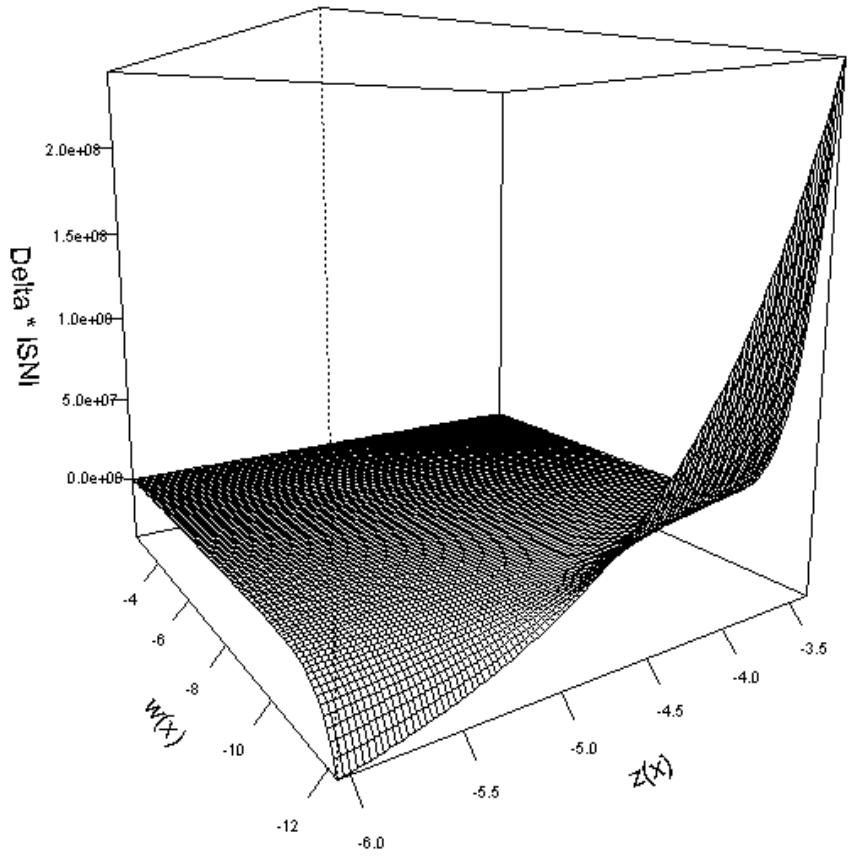

Figure 3.10: Plot showing δ times ISNI, over the range of observed values for $\hat{z}_0(\mathbf{x})$ and $\hat{w}_0(\mathbf{x})$ for patients in the Liver Registration data set, using $\delta = 0.2$

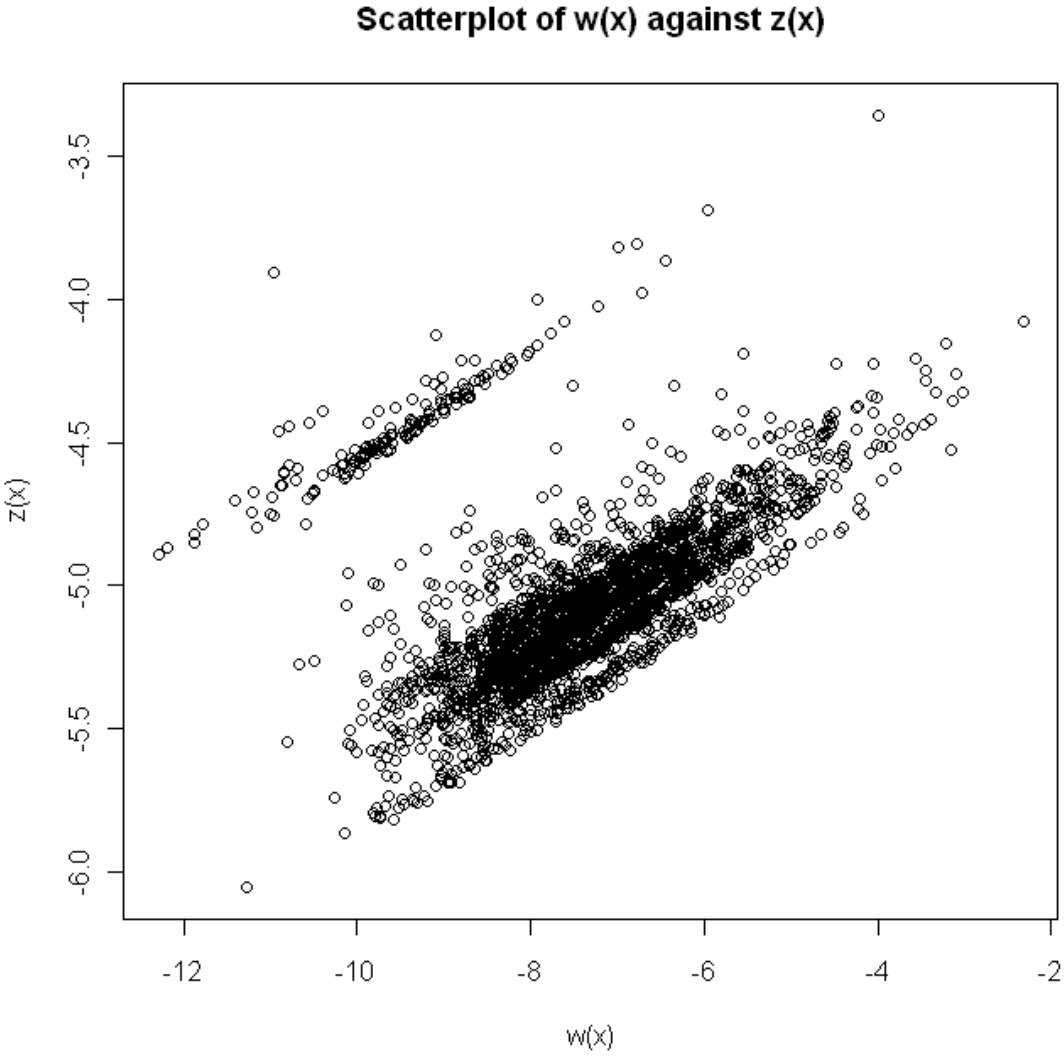

Figure 3.11: Scatterplot showing the observed combinations of $\hat{w}_0(\mathbf{x})$ and $\hat{z}_0(\mathbf{x})$ for all the patients in the Liver Registration data set.

modelled by a Cox proportional hazards model. The dependence between T and C is modelled by an assumed copula function. The parameter of this copula function determines the degree of association that is being assumed. It would be possible to use this model to conduct a sensitivity analysis for the Cox proportional hazards model by varying the parameter over a sensible range.

We assume that the marginal hazard functions for T_i and C_i are

$$h_{Ti}(t|\mathbf{Z}_i, \mathbf{X}_i) = h_{T0}(t)\exp(\mathbf{Z}_i'\boldsymbol{\beta}_T),$$
$$h_{Ci}(t|\mathbf{Z}_i, \mathbf{X}_i) = h_{C0}(t)\exp(\mathbf{X}_i'\boldsymbol{\beta}_C),$$

where $\boldsymbol{\beta}_T$ and $\boldsymbol{\beta}_C$ are unknown parameters, $\mathbf{Z}$ and $\mathbf{X}$ are covariate vectors and $h_{T0}(t)$ and $h_{C0}(t)$ are unspecified baseline hazard functions. Their cumulative hazard functions are denoted by $H_{T0}(t)$ and $H_{C0}(t)$ respectively. We denote their marginal cumulative distributions functions by $F_{Ti}(t)$ and $F_{Ci}(t)$ and survival functions by $S_{Ti}(t)$ and $S_{Ci}(t)$. If we suppose that $\mathcal{C}(u, v; \alpha)$ is a copula with parameter α, then the joint cumulative distribution function of T_i and C_i is given by

$$Pr(T_i \leq t, C_i \leq c) = \mathcal{C}\{F_{Ti}(t), F_{Ci}(c), \alpha\}$$

3.5.1 Fitting an extended Cox model that allows for informative censoring

To develop an extended Cox model that allows for informative censoring, Huang and Zhang (2008) use the idea of "redistribution of mass" that is used in Efron (1967) to derive self-consistent estimators. This idea was briefly explained in Section 2.8, when the self-consistent estimators that use an assumed copula from Zheng and Klein (1994) were reviewed, but will be included in more detail in this section.

Assume that y_i, $i = 1, \ldots, n$, are sorted observation times in ascending order without ties. If y_i is a death time then it is known whether T_i is smaller or greater than t. If y_i is a censored observation time that is greater than or equal to t, then it is also known that the T_i for this individual is greater than t. However, if y_i is a censored observation that is less than t, it is not known if T_i is greater than t as it could fall between y_i and t. Therefore, some assumption needs to be made about the probability that T_i is greater than t.

If the censoring is assumed to be non-informative then it is assumed that a censored individual has equal chance of failure at all event times after their observed censoring time. If there is potentially informative censoring in a dataset, then censored individuals would no longer have equal change of failure at all event times after their observed censoring time. One way of specifying the probability of failure at the event times after their observed censoring time is to use a copula function, as in Zheng and Klein (1994).

Zheng and Klein (1994) show that under the joint distribution assumption specified by the copula $\mathcal{C}$, if the subject i is censored at time y_i, then for each time point $y_j > y_i$, the probability that this subject i fails at time y_j is

$$\begin{aligned} Pr(T_i \geq y_j | T_i \geq y_i, C_i = y_i) &= \frac{Pr(T_i \geq y_j, C_i = y_i)}{Pr(T_i \geq y_i, C_i = y_i)} \\ &= \frac{1 - \mathcal{C}_v\{F_{Ti}(y_j), F_{Ci}(y_i)\}}{1 - \mathcal{C}_v\{F_{Ti}(y_i), F_{Ci}(y_i)\}}, \end{aligned} \tag{3.37}$$

where $\mathcal{C}_v(a,b) = \frac{\partial \mathcal{C}(u,v)}{\partial v}|_{(u,v)=(a,b)}$. We denote the above conditional survival probability by $P_i(y_j)$, so then the mass that subject i loses at time y_j is

$$D_i(y_j) = P_i(y_{j-1}) - P_i(y_j). \tag{3.38}$$

Similarly, all the other subjects censored before time y_j lose some mass at time point y_j. So we define an extended partial likelihood function as follows:

$$L_j^{(T)}(\boldsymbol{\beta}_T) = \prod_{i=1}^{j} \left\{ \frac{P_i(y_j)\exp(\mathbf{Z}_i'\boldsymbol{\beta}_T)}{\sum_{k=1}^{n} P_k(y_j)\exp(\mathbf{Z}_k'\boldsymbol{\beta}_T)} \right\}^{D_i(y_j)}, \tag{3.39}$$

$$\begin{aligned} L^{(T)}(\boldsymbol{\beta}_T) &= \prod_{j=1}^{n} L_j^{(T)}(\boldsymbol{\beta}_T) \\ &= \prod_{j=1}^{n}\prod_{i=1}^{j} \left\{ \frac{P_i(y_j)\exp(\mathbf{Z}_i'\boldsymbol{\beta}_T)}{\sum_{k=1}^{n} P_k(y_j)\exp(\mathbf{Z}_k'\boldsymbol{\beta}_T)} \right\}^{D_i(y_j)}. \end{aligned} \tag{3.40}$$

Here, $L_j^{(T)}(\boldsymbol{\beta}_T)$ is the likelihood function for the time point y_j. So that the above equation is well defined, we need to set $P_k(y_j) = 1$ for $k \geq j$. For a failed subject i, set $P_i(y_j) = 1$ for $j \leq i$, and $P_i(y_j) = 0$ and $j > i$. Also for failed subjects, we do not use (3.38), instead we set $D_i(y_i) = 1$, $D_i(y_j) = 0$ for $j > i$. So a failed subject contributes only one term in this extended partial likelihood function.

When there are tied failure events, then the above equations can naturally handle them using Breslow's method of handling ties. Also, if the pieces of mass $D_i(y_j)$, $i = 1, \ldots, j$, are viewed as the number ties at time y_j, then we would obtain the form of $L_j^{(T)}(\boldsymbol{\beta}_T)$ in (3.39) using Breslow's method.

As well as the extended partial likelihood for the failure events, a similar expression is needed for the censored events. If subject i fails at time y_i, then for $c > y_i$, we have

$$\begin{aligned} Pr(C_i \geq c | C_i \geq y_i, T_i = y_i) &= \frac{Pr(C_i \geq c, T_i = y_i)}{Pr(C_i \geq y_i, T_i = y_i)} \\ &= \frac{1 - \mathcal{C}_u\{F_{Ti}(y_i), F_{Ci}(c)\}}{1 - \mathcal{C}_u\{F_{Ti}(y_i), F_{Ci}(y_i)\}}, \end{aligned} \tag{3.41}$$

where $\mathcal{C}_u(a,b) = \frac{\partial \mathcal{C}(u,v)}{\partial u}|_{(u,v)=(a,b)}$. We denote the above conditional survival probability by $Q_i(c)$, so then the mass that subject i loses at time y_j is

$$U_i(y_j) = Q_i(y_{j-1}) - Q_i(y_j).$$

The extended partial likelihood function for censoring events is given by

$$L^{(C)}(\boldsymbol{\beta}_C) = \prod_{j=1}^{n} \prod_{i=1}^{j} \left\{ \frac{Q_i(y_j) \exp(\mathbf{X}_i' \boldsymbol{\beta}_C)}{\sum_{k=1}^{n} Q_k(y_j) \exp(\mathbf{X}_k' \boldsymbol{\beta}_C)} \right\}^{U_i(y_j)}. \tag{3.42}$$

So that the above equation well defined, we need to set $Q_k(y_j) = 1$ for $k \geq j$. For a censored subject i, set $Q_i(y_j) = 1$ for $j \leq i$, and $Q_i(y_j) = 0$ and $j > i$. Also we set $U_i(y_i) = 1$, $U_i(y_j) = 0$ for $j > i$. For an administratively censored subject i, we set $P_i(y_j) = Q_i(y_j) = 1$ for $j \leq i$, $P_i(y_j) = Q_i(y_j) = 0$ for $j > i$ and $D_i(y_j) = U_i(y_j) = 0$ for all j.

We can now estimate the parameters $\boldsymbol{\beta}_T$ and $\boldsymbol{\beta}_C$ by maximising the following extended joint partial likelihood,

$$L(\boldsymbol{\beta}_T, \boldsymbol{\beta}_C) = L^{(T)}(\boldsymbol{\beta}_T) L^{(C)}(\boldsymbol{\beta}_C). \tag{3.43}$$

Because the functions in this likelihood involve unknown quantities, then we have to carry out the following iterative process:

1. Assuming independent censoring, fit two Cox proportional hazards models to get initial estimators $\hat{\boldsymbol{\beta}}_T^{(0)}$ and $\hat{\boldsymbol{\beta}}_C^{0}$ for $\boldsymbol{\beta}_T$ and $\boldsymbol{\beta}_C$. Then use the Breslow method to obtain the estimators $\hat{H}_{T0}^{(0)}(\cdot)$ for $H_{T0}(\cdot)$ and $\hat{H}_{C0}^{(0)}(\cdot)$ for $H_{C0}(\cdot)$. Let $m = 0$.

2. For $i = 1, \ldots, n$, compute

$$\hat{S}_{Ti}^{(m)}(t) = \exp\left\{ -\hat{H}_{T0}^{(m)}(t) \exp(\mathbf{Z}_i' \hat{\boldsymbol{\beta}}_T^{(m)}) \right\},$$
$$\hat{S}_{Ci}^{(m)}(t) = \exp\left\{ -\hat{H}_{C0}^{(m)}(t) \exp(\mathbf{X}_i' \hat{\boldsymbol{\beta}}_C^{(m)}) \right\}.$$

   Then for each time point y_j such that $y_j > y_i$, compute

$$\hat{P}_i^{(m)}(y_j) = \frac{1 - \mathcal{C}_v\{\hat{F}_{Ti}^{(m)}(y_j), \hat{F}_{Ci}^{(m)}(y_i); \alpha\}}{1 - \mathcal{C}_v\{\hat{F}_{Ti}^{(m)}(y_i), \hat{F}_{Ci}^{(m)}(y_i); \alpha\}},$$
$$\text{if subject } i \text{ is censored;}$$
$$\hat{Q}_i^{(m)}(y_j) = \frac{1 - \mathcal{C}_u\{\hat{F}_{Ti}^{(m)}(y_i), \hat{F}_{Ci}^{(m)}(y_j); \alpha\}}{1 - \mathcal{C}_u\{\hat{F}_{Ti}^{(m)}(y_i), \hat{F}_{Ci}^{(m)}(y_i); \alpha\}},$$
$$\text{if subject } i \text{ is failed;}$$

3. Using the above computation results and other specifications as described earlier, replace the unknown functions P_i, Q_i, D_i, U_i in $L^{(T)}(\boldsymbol{\beta}_T)$ and $L^{(C)}(\boldsymbol{\beta}_C)$ by their estimates at step m, and then maximise the likelihood functions in (3.40) and (3.42) with respect to $\boldsymbol{\beta}_T$ and $\boldsymbol{\beta}_C$, respectively. The resulting estimators for $\boldsymbol{\beta}_T$ and $\boldsymbol{\beta}_C$ are denoted by $\hat{\boldsymbol{\beta}}_T^{(m+1)}$ and $\hat{\boldsymbol{\beta}}_C^{(m+1)}$.

4. Use $\hat{\boldsymbol{\beta}}_T^{(m+1)}$, $\hat{\boldsymbol{\beta}}_C^{(m+1)}$, $\hat{P}_i^{(m)}(\cdot)$, $\hat{Q}_i^{(m)}(\cdot)$, $\hat{D}_i^{(m)}(\cdot)$ and $\hat{U}_i^{(m)}(\cdot)$ to obtain the Breslow estimators $\hat{H}_{T0}^{(m+1)}(\cdot)$ for $H_{T0}(\cdot)$ and $\hat{H}_{C0}^{(m+1)}(\cdot)$ for $H_{C0}(\cdot)$, as shown below:

$$\hat{H}_{T0}^{(m+1)}(t) = \sum_{j:y_j \leq t} \frac{\sum_{i:y_i \leq y_j} \hat{D}_i^{(m)}(y_j)}{\sum_{k=1}^{n} \hat{P}_k^{(m)}(y_j) \exp(\mathbf{Z}_k' \hat{\boldsymbol{\beta}}_T^{(m+1)})},$$

$$\hat{H}_{C0}^{(m+1)}(t) = \sum_{j:y_j \leq t} \frac{\sum_{i:y_i \leq y_j} \hat{U}_i^{(m)}(y_j)}{\sum_{k=1}^{n} \hat{Q}_k^{(m)}(y_j) \exp(\mathbf{X}_k' \hat{\boldsymbol{\beta}}_C^{(m+1)})}.$$

5. Let $m = m + 1$, return to Step 2, and iterate until convergence.

After convergence, we get estimators $\hat{\boldsymbol{\beta}}_T$, $\hat{\boldsymbol{\beta}}_C$, $\hat{H}_{T0}(\cdot)$ and $\hat{H}_{C0}(\cdot)$, respectively, for $\boldsymbol{\beta}_T$, $\boldsymbol{\beta}_C$, $H_{T0}(\cdot)$ and $H_{C0}(\cdot)$.

3.5.2 Applying sensitivity analysis to the Liver Registration dataset

Before using the method presented above to estimate the bivariate proportional hazards model for competing risks, we fitted Cox proportional hazards models for both time to censoring and time to death assuming independent censoring. We selected covariates for these models using a stepwise selection algorithm. We set a p-value of 0.15 as the threshold for variables both to be entered into and stay in the model. This is larger than would usually be used, but this is because p-values will change under the model presented here, and we want to include any variables that might become significant. Also we shall include covariates that are included in the Cox proportionals hazards models for either time to death or time to censoring, so that we are including more covariates than used in other methods.

Table 3.3 compares the results for models for time to death under independent censoring and assuming that the dependence between T and C is modelled by a Clayton copula with Kendall's $\tau = 0.2$.

This table shows that we draw roughly the same conclusions under the two models. All the parameter estimates remain the same sign with the exception those for serum creatinine and patients with AIDS. However, the p-values show that both of these parameter estimates are not significant under either the Cox model or the Huang-Zhang model. There are only a small number of levels of categorical variables that have gone from being non-significant under the model assuming independent censoring to significant under the Huang-Zhang model. These are for patients with AB blood group, patients of black ethnicity, patients with alcoholic liver disease and patients with liver diseases that are not included in any of the other main categories.

We can also examine the changes between the parameter estimates in Table 3.3 to see whether the standard Cox proportional hazards model overestimated or underestimated

	Cox PH model			Huang-Zhang model		
Variable	$\hat{\beta}$	s.e.	p value	$\hat{\beta}$	s.e.	p value
Age	0.0330	0.0063	< 0.001	0.0273	0.0031	< 0.001
Height	0.0039	0.0072	0.586	0.0067	0.0039	0.091
Blood group - AB	0.5012	0.3225	0.120	0.7324	0.1359	< 0.001
Blood group - B	0.4196	0.2013	0.037	0.4314	0.0969	< 0.001
Blood group - O	0.1717	0.1311	0.190	0.0510	0.0647	0.430
Sex - Male	-0.0150	0.1669	0.928	-0.0140	0.0875	0.873
Ethnicity - Black	1.0170	0.5938	0.087	0.5327	0.2666	0.046
Ethnicity - Other	-0.3268	0.7815	0.676	-0.2090	0.3176	0.511
Ethnicity - White	1.0974	0.3200	0.001	0.8004	0.1360	< 0.001
INR	-0.2862	0.1166	0.014	-0.1598	0.0527	0.002
Bilirubin	-0.0010	0.0007	0.191	-0.0002	0.0003	0.643
Sodium	0.0918	0.0257	< 0.001	0.0654	0.0121	< 0.001
UKELD	0.2984	0.0353	< 0.001	0.2374	0.0161	< 0.001
Creatinine	-0.0003	0.0012	0.771	0.0001	0.0007	0.868
PLD - AID	0.0289	0.3481	0.934	-0.0772	0.1690	0.648
PLD - ALD	-0.4126	0.3363	0.220	-0.4593	0.1609	0.004
PLD - Cancer	-1.1880	0.7813	0.128	-0.3808	0.2310	0.099
PLD - HBV	-0.4510	0.5889	0.444	-0.4160	0.2530	0.100
PLD - HCV	0.2904	0.3457	0.401	0.1856	0.1669	0.266
PLD - Metabolic	0.8390	0.3700	0.023	0.7568	0.1848	< 0.001
PLD - Other	0.4295	0.3467	0.216	0.3359	0.1694	0.047
PLD - PBC	-0.2154	0.3634	0.553	-0.2665	0.1741	0.126
PLD - PSC	-1.0029	0.4063	0.014	-0.8413	0.1809	< 0.001

Table 3.3: Results from fitting a standard Cox proportional hazards model and the Huang-Zhang model using a Clayton copula with Kendall's $\tau = 0.2$

the hazard ratios for each covariate. If there is an increase in the value of $\hat{\beta}$ from the Cox model to the Huang-Zhang model, then the estimated hazard ratio would also increase. This suggests that the hazard ratio is underestimated by the standard Cox model. Conversely, if there is a decrease in $\hat{\beta}$ from the standard Cox model to the Huang-Zhang model then the hazard ratio is being overestimated by the standard Cox model. The results in Table 3.3 suggest that the standard Cox model overestimates the hazard ratio for the following covariates and levels of factors: height, blood groups AB and B, males, ethnicities that are not one of the main categories, INR, serum bilirubin, and patients with cancer, hepatitis B virus or primary sclerosing cholangitis. The parameter estimates in Table 3.3 also suggest that the standard Cox model underestimates the hazard ratio for the following covariates and levels of factors: age, blood group O, black and white ethnicity, serum sodium, UKELD score and patients with alcoholic liver disease, hepatitis C virus, metabolic liver disease, primary biliary cirrhosis or other liver diseases not included in the main categories.

3.6 Siannis (2011) Sensitivity Analysis

The sensitivity analysis here uses a similar approach to that of Siannis (2004) and Siannis et al. (2005) but instead of considering parametric survival models, it uses the Cox proportional hazards model.

It is still assumed that

$$f_{C|T}(c|t,\gamma,\delta,\theta) = f_C(c;\gamma + \delta i_\gamma^{-1/2} B(t,\theta)), \tag{3.44}$$

but now this function is written in terms of an unspecified baseline hazard multiplied by a parametric function instead of a known parametric baseline hazard. The Cox proportional hazards model assumes that

$$h_T(t,\theta_i) = e^{\theta_i} h_{T0}(t),$$

with θ_i usually expressed as $\boldsymbol{\beta}_T' \mathbf{x}_i$ to incorporate the covariates in the vector $\mathbf{x}_i$. Therefore the quantity of interest is now $\boldsymbol{\beta}_T$. The hazard function for C can be expressed in a similar form, with the corresponding vector of regression coefficients denoted by $\boldsymbol{\beta}_C$. For simplicity, C is assumed to have the same covariate vector $\mathbf{x}_i$ as T.

To derive a partial likelihood for the proportional hazards model when there is potentially informative censoring, the competing risks set up is used. This means that T and C are seen as two competing causes of failure with the observed time for individual i being $Y_i = \min(T_i, C_i)$. The cause of the failure, J, is also observed. As there are two competing causes of failure, then J can take values 1 or 2. Let $J = 1$ denote that the observed time is a failure time and $J = 2$ denote that it is a censored observation. The competing risks

set up considers the sub-hazard function of cause j, which is defined as

$$h(j,t,\theta) = \lim_{dt\to 0^+} \frac{P(t \le Y < t+dt, J = j | Y \ge t, \theta)}{dt}$$

and is the hazard of failure from cause j in the presence of all the other causes. The sub-hazard function $h(T,t,\theta)$ will be the same as the marginal hazard function $h_T(t,\theta)$ if there is no informative censoring.

The competing risks partial likelihood,

$$L_p^* = \prod_{i=1}^{n} \frac{h(j,t_i|\mathbf{x}_i)}{\sum_{l\in\mathcal{R}_{t_i}} h(j,t_l|\mathbf{x}_l)},$$

where $\mathcal{R}_{t_i}$ is the risk set at time t_i, uses sub-hazard functions instead of marginal hazard functions like the ordinary partial likelihood. When the two causes being considered are $T$and C, each death time contributes

$$\frac{h(T,t|\mathbf{x}_i)}{\sum_{l\in\mathcal{R}_t} h(T,t|\mathbf{x}_i)}$$

to the likelihood, while each censored time c contributes

$$\frac{h(C,c|\mathbf{x}_i)}{\sum_{q\in\mathcal{R}_c} h(C,c|\mathbf{x}_i)}$$

to the likelihood. The product of the contributions from all the individuals in the data set gives the modified partial likelihood (MPL),

$$L_M = \prod_{i=1}^{n_T} \frac{h(T,t_i|\mathbf{x}_i)}{\sum_{l\in\mathcal{R}_{t_i}} h(T,t_l|\mathbf{x}_l)} \prod_{k=1}^{n_C} \frac{h(C,t_k|\mathbf{x}_k)}{\sum_{q\in\mathcal{R}_{t_k}} h(C,t_q|\mathbf{x}_q)} \tag{3.45}$$

where n_T is the number of deaths and n_C is the number of censored observations. If the censoring is assumed to be ignorable, then the sub-hazards in (3.45) would be equal to the marginal hazards and it would become

$$L_p = \prod_{i=1}^{n_T} \frac{e^{\boldsymbol{\beta}_T'\mathbf{x}_i}}{\sum_{l\in\mathcal{R}_{t_i}} e^{\boldsymbol{\beta}_T'\mathbf{x}_l}} \prod_{k=1}^{n_C} \frac{e^{\boldsymbol{\beta}_C'\mathbf{x}_k}}{\sum_{q\in\mathcal{R}_{t_k}} e^{\boldsymbol{\beta}_C'\mathbf{x}_q}},$$

which is the product of two ordinary partial likelihoods.

Using the assumption in (3.44), Siannis (2011) show that a first-order approximation with respect to δ of the sub-hazard function of T is

$$h(T,t|\mathbf{x}) \simeq h_T(t,\boldsymbol{\beta}_T|\mathbf{x}) \left[1 + \delta i_\gamma^{-1/2} \mu_T(t,\boldsymbol{\beta}_T|\mathbf{x})\psi(t|\mathbf{x})\right] \tag{3.46}$$

where

$$\mu(t,\boldsymbol{\beta}_T|\mathbf{x}) = \frac{\int_t^\infty B(u,\boldsymbol{\beta}_T|\mathbf{x}) f_T(u,\boldsymbol{\beta}_T|\mathbf{x})du}{S_T(t,\boldsymbol{\beta}_T|\mathbf{x})},$$
$$\mu_T(t,\boldsymbol{\beta}_T|\mathbf{x}) = \frac{\partial\mu(t,\boldsymbol{\beta}_T|\mathbf{x})}{\partial T} \quad \text{and} \quad \psi(t|\mathbf{x}) = \frac{H_C(t,\boldsymbol{\beta}_C|\mathbf{x})}{h_T(t,\boldsymbol{\beta}_T|\mathbf{x})}.$$

Similarly it can be shown that an approximation of the sub-hazard function for C is

$$h(C,t|\mathbf{x}) \simeq h_C(t,\boldsymbol{\beta}_C|\mathbf{x})\left[1+\delta i_\gamma^{-1/2}\mu(t,\boldsymbol{\beta}_T|\mathbf{x})\right]. \tag{3.47}$$

The approximations in (3.46) and (3.47) can be substituted into (3.45) so that the MPL becomes

$$L_M = \prod_{i=1}^{n_T}\left\{\frac{e^{\boldsymbol{\beta}_T'\mathbf{x}_i}\left[1+\delta i_\gamma^{-1/2}\mu_T(t_i,\boldsymbol{\beta}_T|\mathbf{x}_i)\psi(t_i|\mathbf{x}_i)\right]}{\sum_{l\in\mathcal{R}_{t_i}} e^{\boldsymbol{\beta}_T'\mathbf{x}_l}\left[1+\delta i_\gamma^{-1/2}\mu_T(t_l,\boldsymbol{\beta}_T|\mathbf{x}_l)\psi(t_l|\mathbf{x}_l)\right]}\right\}$$
$$\times\prod_{k=1}^{n_C}\left\{\frac{e^{\boldsymbol{\beta}_C'\mathbf{x}_k}\left[1+\delta i_\gamma^{-1/2}\mu(t_k,\boldsymbol{\beta}_T|\mathbf{x}_k)\right]}{\sum_{q\in\mathcal{R}_{t_k}} e^{\boldsymbol{\beta}_C'\mathbf{x}_q}\left[1+\delta i_\gamma^{-1/2}\mu(t_q,\boldsymbol{\beta}_T|\mathbf{x}_q)\right]}\right\}. \tag{3.48}$$

The MPL in (3.48) can be manipulated to obtain an approximation of the difference between the estimated regression coefficients for T under the assumption of informative censoring and under the assumption of non-informative censoring,

$$\hat{\boldsymbol{\beta}}_{T\delta}-\hat{\boldsymbol{\beta}}_{T0} = \delta\{i(\boldsymbol{\beta}_T)\}^{-1}\{\sum_{i=1}^{n_T}[\frac{\partial G_i}{\partial\boldsymbol{\beta}_T}]+\sum_{k=1}^{n_C}[\frac{\partial K_k}{\partial\boldsymbol{\beta}_T}-\mathbf{x}_k H_T(t_k,\boldsymbol{\beta}_T|\mathbf{x}_k)]\}, \tag{3.49}$$

where

$$G_i = \frac{\sum_{l\in\mathcal{R}_{t_i}} e^{\boldsymbol{\beta}_T'\mathbf{x}_l}H_C(t_i,\boldsymbol{\beta}_C|\mathbf{x}_l)}{\sum_{l\in\mathcal{R}_{t_i}} e^{\boldsymbol{\beta}_T'\mathbf{x}_l}}$$
$$K_k = \frac{\sum_{q\in\mathcal{R}_{t_k}} e^{\boldsymbol{\beta}_C'\mathbf{x}_q}H_T(t_k,\boldsymbol{\beta}_T|\mathbf{x}_q)}{\sum_{q\in\mathcal{R}_{t_k}} e^{\boldsymbol{\beta}_C'\mathbf{x}_q}} \quad \text{and}$$
$$i(\boldsymbol{\beta}_T) = -\frac{\partial^2\log L_p}{\partial\boldsymbol{\beta}_T\partial\boldsymbol{\beta}_T'}.$$

Equation (3.49) can be used to conduct a sensitivity analysis for $\boldsymbol{\beta}_T$ as δ makes small departures from zero. This involves much more computation than the sensitivity analysis presented by Siannis (2004) and Siannis et al. (2005). Estimation of the baseline hazard function is required for use in the cumulative hazard functions in (3.49).

This sensitivity analysis can be extended to the situation where there are several types of censoring, one of which is potentially informative and one which is ignorable. The informative censoring process, C_I, is allowed to contribute information to the likelihood as before and the ignorable censoring process C_E contributes only to the definition of the risk sets. So if there are n_{C_I} potentially informative censored observations out of a total of n_C censored observations, then the MPL takes the form,

$$L_M^I = \prod_{i=1}^{n_T}\frac{h(T,t_i|\mathbf{x}_i)}{\sum_{l\in\mathcal{R}_{t_i}}h(T,t_i|\mathbf{x}_i)}\prod_{k=1}^{n_{C_I}}\frac{h(C_I,t_k|\mathbf{x}_k)}{\sum_{q\in\mathcal{R}_{t_k}}h(C_I,t_q|\mathbf{x}_q)}.$$

This is similar to (3.45) but with the second product only over the n_{C_I} potentially informatively censored observations, rather than all the censored observations.

3.6.1 Application to the Liver Registration data set

The sensitivity analysis for the Cox proportional hazards model that was described in Section 3.6 is applied to the Liver Registration data set in this section. As both potentially informative and non-informative censoring is observed in this data set, then the sensitivity analysis that allows several types of censoring is used. This means the sensitivity analysis equation in (3.49) becomes

$$\hat{\boldsymbol{\beta}}_{T\delta} - \hat{\boldsymbol{\beta}}_{T0} = \delta\{i(\boldsymbol{\beta}_T)\}^{-1}\{\sum_{i=1}^{n_T}[\frac{\partial G_i}{\partial \boldsymbol{\beta}_T}] + \sum_{k=1}^{n_{C_I}}[\frac{\partial K_k}{\partial \boldsymbol{\beta}_T} - \mathbf{x}_k H_T(t_k, \boldsymbol{\beta}_T|\mathbf{x}_k)]\}. \tag{3.50}$$

The equation in (3.50) is almost identical to (3.49), but with the second summation only over those patients who are potentially informatively censored. We see that to apply the sensitivity analysis, we require estimates of β_T and β_C, as well as the baseline hazard functions $H_T(t_k, \boldsymbol{\beta}_T|\mathbf{x}_k)$ and $H_C(t_k, \boldsymbol{\beta}_C|\mathbf{x}_k)$.

The parameter estimates substituted into (3.50) will be those from the Cox proportional hazards model assuming non-informative censoring. The estimated values of β_T and β_C when it is assumed that $\delta = 0$ are given in Tables 3.4 and 3.5. The same covariates were used in both the model for time to failure and the model for time to censoring. These covariates were primary liver disease category, ethnicity, UKELD score, age, serum sodium at time of registration and INR at time of registration. These are the variables that were found to be significant for time to failure when fitting a Cox proportional hazards model.

The baseline hazard functions in (3.50) are estimated by the Breslow estimate of the baseline cumulative hazard function. This is a step function where

$$\tilde{H}_0(t) = \sum_{j=1}^{k} \frac{d_j}{\sum_{l \in R(t_{(j)})} \exp(\hat{\boldsymbol{\beta}}' \mathbf{x}_l)}$$

for $t_{(k)} \leq t < t_{(k+1)}$, $k = 1, 2, \ldots, r-1$, where d_j is the number of events at the jth ordered event time $t_{(j)}$ and r is the total number of events observed.

Table 3.6 shows the estimated values of the components of $\hat{\boldsymbol{\beta}}_{T\delta} - \hat{\boldsymbol{\beta}}_{T0}$ for $\delta = 0.2$ and $\delta = 0.3$. Positive changes in the parameter estimates mean that the hazard ratios of the covariates are being underestimated by the model assuming non-informative censoring. Conversely, negative changes in the parameter estimates mean that the hazard ratios for the covariates are overestimated by the model that assumes $\delta = 0$.

From Table 3.6, we can see that the sensitivity analysis for the Cox proportional hazards model suggests most of the hazard ratios for the levels of primary liver disease category are being overestimated by the model that assumes non-informative censoring, with the exception of patients with cancer or hepatitis B infection. The sensitivity analysis also suggests that the hazard ratios for all levels of patient ethnicity, the UKELD score,

Parameter	Estimate	Standard Error	p-value	Hazard Ratio	95% Confidence Interval for Parameter Estimate
PLD - PBC	-0.1647	0.3442	0.632	0.848	(-0.839,0.510)
PLD - PSC	-0.9295	0.3978	0.020	0.395	(-1.709,-0.150)
PLD - ALD	-0.2665	0.3222	0.408	0.766	(-0.898,0.365)
PLD - AID	0.1275	0.3392	0.707	1.136	(-0.537,0.792)
PLD - HCV	0.4138	0.3371	0.220	1.513	(-0.247,1.075)
PLD - HBV	-0.4116	0.5820	0.480	0.663	(-1.552,0.729)
PLD - Cancer	-1.0664	0.7754	0.169	0.344	(-2.586,0.453)
PLD - Metabolic	0.9208	0.3662	0.012	2.511	(0.203,1.639)
PLD - Other	0.4064	0.3438	0.237	1.501	(-0.267,1.080)
PLD - Acute	0				
Ethnicity - White	1.1246	1.0080	0.265	3.079	(-0.851,3.100)
Ethnicity - Asian	0.0538	1.0484	0.959	1.055	(-2.001,2.109)
Ethnicity - Black	1.0239	1.1218	0.361	2.784	(-1.175,3.223)
Ethnicity - Chinese	-0.4562	1.4345	0.751	0.634	(-3.268,2.355)
Ethnicity - Other	0				
UKELD	0.2628	0.0192	< 0.001	1.301	(0.225,0.300)
Age	0.0331	0.0062	< 0.001	1.034	(0.021,0.045)
Sodium	0.0701	0.0169	< 0.001	1.073	(0.037,0.103)
INR	-0.2300	0.0986	0.020	0.795	(-0.423,-0.037)

Table 3.4: The parameter estimates, estimated hazard ratios, p-values and 95% confidence intervals from the Cox model for time to failure assuming non-informative censoring

Parameter	Estimate	Standard Error	p-value	Hazard Ratio	95% Confidence Interval for Parameter Estimate
PLD - PBC	-0.2282	0.1448	0.115	0.796	(-0.512,0.056)
PLD - PSC	-0.1403	0.1447	0.333	0.869	(-0.424,0.143)
PLD - ALD	-0.2081	0.1353	0.124	0.812	(-0.473,0.057)
PLD - AID	-0.1788	0.1448	0.217	0.836	(-0.463,0.105)
PLD - HCV	0.0545	0.1400	0.697	1.056	(-0.220,0.329)
PLD - HBV	0.1567	0.1815	0.388	1.170	(-0.199,0.512)
PLD - Cancer	0.6728	0.1585	< 0.001	1.960	(0.362,0.983)
PLD - Metabolic	0.0645	0.1705	0.705	1.067	(-0.270,0.399)
PLD - Other	-0.4106	0.1492	0.006	0.663	(-0.703,-0.118)
PLD - Acute	0				
Ethnicity - White	0.3841	0.2550	0.132	1.468	(-0.116,0.884)
Ethnicity - Asian	0.3081	0.2635	0.242	1.361	(-0.208,0.825)
Ethnicity - Black	0.4820	0.2959	0.103	1.619	(-0.098,1.062)
Ethnicity - Chinese	0.8348	0.3731	0.025	2.304	(0.104,1.566)
Ethnicity - Other	0				
UKELD	0.0538	0.0078	< 0.001	1.055	(0.038,0.069)
Age	0.0012	0.0023	0.595	1.001	(-0.003,0.006)
Sodium	0.0316	0.0076	< 0.001	1.032	(0.017,0.046)
INR	-0.0405	0.0416	0.330	0.960	(-0.122,0.041)

Table 3.5: The parameter estimates, estimated hazard ratios, p-values and 95% confidence intervals from the Cox model for time to censoring assuming non-informative censoring

serum sodium and INR are being underestimated by the model that assumes $\delta = 0$. Finally, the sensitivity analysis suggests that the hazard ratio for age is being slightly overestimated.

The results of this sensitivity analysis for the Cox proportional hazards model in Table 3.6 can be compared to the results of the sensitivity analysis for a Weibull proportional hazards model in Table 3.1 in Section 3.3.2. Both of the sensitivity analyses are approximating the change in parameter estimates if informative censoring is assumed instead of non-informative censoring. Both models include primary liver disease category, age, ethnicity and UKELD score, so the estimated changes in the parameter estimates for these variables can be compared.

The sensitivity analyses applied in this section and Section 3.3.2 give similar results for the majority of the parameter estimates. However, the sensitivity analysis from Siannis (2004) suggests that the hazard ratio for cancer patients is overestimated while the sensitivity analysis from Siannis (2011) suggest this hazard ratio is being underestimated. Also, the sensitivity analysis from Siannis (2011) suggests that the hazard ratios for all levels of ethnicity are being underestimated, but the sensitivity analysis from Siannis (2004) suggests that the hazard ratios for white and black patients are being overestimated. The two approaches also disagree about the effect of informative censoring on the hazard ratio for the UKELD score. Siannis (2004) suggests it is being overestimated by the model that assumes $\delta = 0$, whereas Siannis (2011) suggests it is underestimated.

Tables 3.7 and 3.8 show the approximate parameter estimates for Cox proportional hazards models assuming $\delta = 0.2$ and $\delta = 0.3$ respectively. These parameter estimates are obtained by adding the parameter estimates from the model assuming non-informative censoring in Table 3.4 to the values in Table 3.6. The p-values of the estimates are also shown. These are all calculated using the standard errors of the estimates from the model assuming non-informative censoring. The reason that we can do this was discussed previously in Section 3.3.2. It was shown in Siannis et al. (2005) that

$$\mathrm{SE}(\hat{\theta}_\delta) \simeq \mathrm{SE}(\hat{\theta}_0) + O(\delta^2).$$

As only linear values of δ are considered in the sensitivity analysis then the standard error of the parameter estimate from the model assuming informative censoring can be approximated by the standard error of the parameter estimate from the model assuming non-informative censoring. Again, this only applies if the value of δ is fairly small.

3.7 Rotnitzky et al. (2007) Sensitivity Analysis

In sections 3.3 to 3.6, the sensitivity analyses assess the sensitivity of the results under the assumption of non-informative censoring if informative censoring is assumed instead. The

Parameter	$\hat{\boldsymbol{\beta}}_{T0.2} - \hat{\boldsymbol{\beta}}_{T0}$	$\hat{\boldsymbol{\beta}}_{T0.3} - \hat{\boldsymbol{\beta}}_{T0}$
PLD - PBC	-0.0625	-0.0937
PLD - PSC	-0.0746	-0.1119
PLD - ALD	-0.0483	-0.0724
PLD - AID	-0.0604	-0.0906
PLD - HCV	-0.0243	-0.0365
PLD - HBV	0.0931	0.1397
PLD - Cancer	0.0172	0.0258
PLD - Metabolic	-0.0423	-0.0634
PLD - Other	-0.0919	-0.1379
Ethnicity - White	0.0502	0.0754
Ethnicity - Asian	0.0605	0.0908
Ethnicity - Black	0.0505	0.0757
Ethnicity - Chinese	0.2376	0.3564
UKELD	0.0039	0.0058
Age	-0.00003	-0.00005
Sodium	0.0028	0.0042
INR	0.0118	0.0177

Table 3.6: The estimated values of $\hat{\boldsymbol{\beta}}_{T\delta} - \hat{\boldsymbol{\beta}}_{T0}$ for $\delta = 0.2$ and $\delta = 0.3$, calculated by applying the Siannis (2011) sensitivity analysis to the Liver Registration data set

Parameter	Estimate	Hazard Ratio	p-value	95% Confidence Interval
PLD - PBC	-0.2272	0.797	0.509	(-0.902,0.573)
PLD - PSC	-1.0041	0.366	0.012	(-1.784,-0.075)
PLD - ALD	-0.3148	0.730	0.329	(-0.946,0.413)
PLD - AID	0.0671	1.069	0.843	(-0.598,0.853)
PLD - HCV	0.3895	1.476	0.248	(-0.271,1.099)
PLD - HBV	-0.3184	0.727	0.584	(-1.646,0.822)
PLD - Cancer	-1.0493	0.350	0.176	(-2.603,0.470)
PLD - Metabolic	0.8786	2.408	0.016	(0.161,1.681)
PLD - Other	0.3145	1.370	0.360	(-0.359,1.172)
PLD - Acute	0			
Ethnicity - White	1.1749	3.238	0.244	(-0.901,3.151)
Ethnicity - Asian	0.1143	1.121	0.913	(-2.062,2.169)
Ethnicity - Black	1.0744	2.928	0.338	(-1.225,3.273)
Ethnicity - Chinese	-0.2186	0.804	0.879	(-3.506,2.593)
Ethnicity - Other	0			
UKELD	0.2667	1.306	< 0.001	(0.221,0.304)
Age	0.0331	1.034	< 0.001	(0.021,0.045)
Sodium	0.0729	1.076	< 0.001	(0.034,0.106)
INR	-0.2300	0.804	0.027	(-0.435,-0.025)

Table 3.7: The parameter estimates, estimated hazard ratios, p-values and 95% confidence intervals for the Cox model assuming informative censoring with $\delta = 0.2$. These values were found using the estimates from the Cox model assuming non-informative censoring and the estimated differences from the Siannis (2011) sensitivity analysis.

Parameter	Estimate	Hazard Ratio	p-value	95% Confidence Interval
PLD - PBC	-0.2584	0.772	0.453	(-0.933,0.604)
PLD - PSC	-1.0414	0.353	0.009	(-1.821,-0.038)
PLD - ALD	-0.3389	0.713	0.293	(-0.970,0.437)
PLD - AID	0.0369	1.038	0.914	(-0.628,0.883)
PLD - HCV	0.3774	1.458	0.263	(-0.283,1.111)
PLD - HBV	-0.2719	0.762	0.640	(-1.692,0.869)
PLD - Cancer	-1.0407	0.353	0.180	(-2.612,0.479)
PLD - Metabolic	0.8575	2.357	0.019	(0.140,1.702)
PLD - Other	0.2685	1.308	0.435	(-0.405,1.218)
PLD - Acute	0			
Ethnicity - White	1.2000	3.320	0.234	(-0.927,3.176)
Ethnicity - Asian	0.1446	1.156	0.890	(-2.092,2.200)
Ethnicity - Black	1.0996	3.003	0.327	(-1.251,3.298)
Ethnicity - Chinese	-0.0998	0.905	0.945	(-3.624,2.712)
Ethnicity - Other	0			
UKELD	0.2686	1.308	< 0.001	(0.219,0.306)
Age	0.0331	1.034	< 0.001	(0.021,0.045)
Sodium	0.0743	1.077	< 0.001	(0.033,0.108)
INR	-0.2123	0.809	0.031	(-0.441,-0.019)

Table 3.8: The parameter estimates, estimated hazard ratios, p-values and 95% confidence intervals for the Cox model assuming informative censoring with $\delta = 0.3$. These values were found using the estimates from the Cox model assuming non-informative censoring and the estimated differences from the Siannis (2011) sensitivity analysis.

sensitivity analysis presented in this section has a different aim from all the sensitivity analyses considered so far. Rotnitzky et al. (2007) derive a sensitivity analysis that assesses the sensitivity of an estimator that accounts for informative censoring by adjusting for measured prognostic factors to different levels of residual dependence. The estimator accounts for potentially informative censoring by assuming sequential ignorability of censoring which means that after adjusting for all the measured prognostic factors, the time to event variables and time to censoring variables are independent of each other. However, it is possible that some of the dependence between the two variables could be due to unmeasured factors, which is called residual dependence.

A semi-parametric model was used in Scharfstein and Robins (2002) to allow for residual dependence between the two variables after incorporating a vector of covariates. However, this model only allowed one censoring mechanism. This meant that either all the censoring in a data set would have to be treated as informative, even if is administrative censoring, or any data after the first occurrence of administrative censoring is disregarded. Scharfstein and Robins (2002) adopt the latter strategy. An extension of this model was presented in Rotnitzky et al. (2007) that allowed for multiple causes of censoring. This is the model that will be used here as there is administrative censoring as well as possibly informative censoring in the Liver Registration data set.

In Section 3.7.1 all the necessary notation for the model is presented and the form of the model that will be used is given. This model will then be used to conduct a sensitivity analysis for the assumption of sequential ignorability of censoring for the Liver Registration data set.

3.7.1 Notation

All of the variables defined in this section come from Rotnitzky et al. (2007), although in some cases the notation used has been changed slightly so that it is consistent with the notation used in the rest of this chapter.

Let T^* and C^* be the times from entry into the study to the time of death and time of censoring, respectively. The maximum follow-up time that will be used when applying this method is κ. As it is possible that either of these events could occur after time κ, we define $T = \min(T^*, \kappa)$ and $C = \min(C^*, \kappa)$. However, $Y = \min(T, C)$ is actually observed.

The maximum possible follow-up time for any patient is κ^*. However, for a technical reason, any data that was recorded after time $\kappa = \kappa^* - \epsilon$, where ϵ is a small positive number, needs to be disregarded. The technical reason is that the condition

$$h_{C,j}(u|V^H(u), T, T > u) < K \quad \text{with probability 1,} \tag{3.51}$$

where

$$h_{C,j}(u|\cdot, T > u) = \lim_{du \to 0} \{Pr(u \leq C < u + du, J = j | C \geq u, T > u, \cdot)/du\},$$

has to hold for all $u \in (0, \kappa)$ and some constant K. Condition (3.51) will be false when $\kappa = \kappa^*$ since all patients who are uncensored just before κ^* will be censored when the study ends at this time. Therefore $\kappa = \kappa^* - \epsilon$ for some $\epsilon > 0$ is used instead.

A vector of covariates $\mathbf{V}(t)$ is recorded at either predetermined or random times. The history of this covariate vector is defined as $\bar{\mathbf{V}}(t) = \{\mathbf{V}(u); 0 \leq u \leq t\}$. The vector of baseline covariates is $\mathbf{V}(0)$, and for $u > 0$, $\mathbf{V}(u)$ is the vector of covariate values at time u, if it happened to be a measurement time, or the last values recorded before u if it was not a measurement time.

As we are allowing several types of censoring within the model, a variable that distinguishes between the different types is needed. Let $J \in \{1, 2, \ldots, j^*\}$ denote the cause of censoring, where j^* is the number of different censoring types. There is also the event indicator $\Delta = I(T \leq C)$, and we let $J = 0$ if $\Delta = 1$.

So, the observed data are the independent and identically distributed $O_1, \ldots, O_n$, where $O_i = (\Delta_i, Y_i, J_i, \bar{\mathbf{V}}(Y_i))$. These will be used to estimate $S_T(t^*) = Pr(T > t^*)$ for any $t^* \in (0, \kappa)$.

We will consider estimators of $S_T(t^*)$ under the following assumption about the censoring variables, for $j = 1, \ldots, j^*$ and $u \in [0, \kappa)$,

$$h_{C,j}\{u|\bar{\mathbf{V}}(u), T, T > u\} = h_{0,j}\{u, \bar{\mathbf{V}}(u)\} \exp[q_j\{u, \bar{\mathbf{V}}(u), T\}] \tag{3.52}$$

where $h_{0,j}\{u, \bar{\mathbf{V}}(u)\}$ is an unknown non-negative function of both u and $\bar{\mathbf{V}}(u)$. The functions $q_j\{u, \bar{\mathbf{V}}(u), T\}$, are known functions of u, $\bar{\mathbf{V}}(u)$ and T, that are called cause-specific censoring bias functions. They measure the dependence on the hazard ratio scale between T and censoring due to cause j at time u, after adjusting for the measured prognostic factors in $\bar{\mathbf{V}}(u)$. If q_j is set to zero, then for censoring cause j, sequential ignorability of censoring is being assumed. This is the assumption used in Robins and Finkelstein (2000) when constructing IPCW estimators. This assumes that time to death and time to censoring are independent after adjusting for covariates that are prognostic factors for both the time to death and time to censoring variables. If no prognostic factors are included in the model and q_j is set to zero, then this is equivalent to the assumption of non-informative censoring.

The model presented in (3.52) is referred to as model A_q. Is is only possible estimate $S_T(t^*)$ under this model when $\bar{\mathbf{V}}(u)$ is low dimensional, such as equal to a single baseline discrete covariate V for all u. This is because estimation of $S_T(t^*)$ under this model

requires estimation of

$$H_{0,j}\{u, \bar{\mathbf{V}}(u)\} = \int_0^u h_{0,j}\{s, \bar{\mathbf{V}}(s)\}\mathrm{d}s, \quad \text{for } j = 1, \ldots, j^*.$$

This is not feasible when $\bar{\mathbf{V}}(u)$ is high dimensional, so when this is the case the dimensionality of the unknown functions $h_{0,j}\{u, \bar{\mathbf{V}}(u)\}$ needs to be reduced. This can be done by assuming a semi-parametric model of the form

$$h_{0,j}\{u, \bar{\mathbf{V}}(u)\} = h_{0,j}(u) \exp[\boldsymbol{\nu}_j' w_j\{u, \bar{\mathbf{V}}(u)\}] \quad j = 1, \ldots, j^* \tag{3.53}$$

where $h_{0,j}(\cdot)$ is an unknown function, $w_j\{u, \bar{\mathbf{V}}(u)\}$ is a specified vector function of $\bar{\mathbf{V}}(u)$ and $\boldsymbol{\nu}_j$ is a vector of unknown parameters.

When the additional restriction in (3.53) is imposed on model A_q, then the resulting model is called model B_q.

3.7.2 Estimation identities

This section outlines the identities that are required to construct the estimator of $S_T(t^*)$, all of which are given in Rotnitzky et al. (2007). The fundamental identity used is

$$E\left[\Delta \frac{\pi\{u|\bar{\mathbf{V}}(u), T; H_0\}}{\pi\{T|\bar{\mathbf{V}}(u), T; H_0\}} \Big| \bar{\mathbf{V}}(u), T, T > u, Y \geq u\right] = 1 \quad \text{for all } u \in (0, \kappa]. \tag{3.54}$$

where

$$\begin{aligned}\pi\{u|\bar{\mathbf{V}}(u), T; H_0\} &= \exp\left[-\int_0^t h_C\{u|\bar{\mathbf{V}}(u), T, T > u\}\mathrm{d}u\right] \\ &= \prod_{j=1}^{j^*} \prod_{0 \leq u \leq t} \left(1 - \exp[q_j\{u, \bar{\mathbf{V}}(u), T\}] dH_{0,j}\{u, \bar{\mathbf{V}}(u)\}\right).\end{aligned}$$

This identity implies that $S_T(t^*)$ is the solution of the population moment equation

$$E\left[\frac{\Delta}{\pi\{T|\bar{\mathbf{V}}(T), T; h_0\}}\{I(T > t^*) - S_T(t^*)\}\right] = 0. \tag{3.55}$$

We shall impose the condition that for all $u \in (0, \kappa)$ and some constant K,

$$h_{C,j}\{u|\bar{\mathbf{V}}(u), T, T > u\} < K \quad \text{with probability 1}$$

so that under this condition and model B_q, (3.54) implies that $H_{0,j}\{u, \bar{\mathbf{V}}(u)\}$ satisfies

$$H_{0,j}(u) = \int_0^u \frac{h_{C,j}(s|T > s)\mathrm{d}s}{E\left(\Delta \frac{\pi\{s|\bar{\mathbf{V}}(s), T; H_0\}}{\pi\{T|\bar{\mathbf{V}}(T), T; H_0\}} \exp\left[\boldsymbol{\nu}_j' w_j\{s, \bar{\mathbf{V}}(s)\} + q_j\{s, \bar{\mathbf{V}}(s), T\}\right] \Big| C \geq s, T > s\right)}. \tag{3.56}$$

It also follows from identity (3.54) that, under model B_q, when $\boldsymbol{\nu}$ and H_0 are set at their true values,

$$E\{\mathrm{d}M_{C,j}(u; H_0, \boldsymbol{\nu})|\bar{\mathbf{V}}(u), T, T > u, Y \geq u\} = 0$$

where

$$\mathrm{d}M_{C,j}(u; H_0, \boldsymbol{\nu}) = \mathrm{d}N_{C,j}(u) - \Delta \frac{\pi\{u|\bar{\mathbf{V}}(u), T; \boldsymbol{\nu}, H_0\}}{\pi\{T|\bar{\mathbf{V}}(T), T; \boldsymbol{\nu}, H_0\}} \mathrm{d}H_{0,j}(u) r_j\{u, \bar{\mathbf{V}}(u), T; \boldsymbol{\nu}_j\}, \quad (3.57)$$

$$r_j\{u, \bar{\mathbf{V}}(u), T; \boldsymbol{\nu}_j\} = \exp\left[\boldsymbol{\nu}_j' w_j\{u, \bar{\mathbf{V}}(u)\} + q_j\{u, \bar{\mathbf{V}}(u), T\}\right]$$

and

$$N_{C,j}(u) = I(C \leq u, J = j).$$

The functions $\pi\{\cdot|\bar{\mathbf{V}}(\cdot), T; \boldsymbol{\nu}, H_0\}$ in (3.57) are defined like $\pi\{\cdot|\bar{\mathbf{V}}(\cdot), T; H_0\}$ but with $\exp[q_j\{u, \bar{\mathbf{V}}(u), T\}]$ and $H_0\{u, \bar{\mathbf{V}}(u)\}$ replaced by $r_j\{u, \bar{\mathbf{V}}(u), T; \boldsymbol{\nu}_j\}$ and $H_0(u)$ respectively.

However, we shall construct estimating equations from $E\{m(O; \boldsymbol{\nu}, H_0, a)\} = 0$, where for any collections of functions $a = \{a_j\{\cdot, \bar{\mathbf{V}}(\cdot)\} : j = 1, \ldots, j^*\}$, we have

$$m(O; \boldsymbol{\nu}, H, a) = \sum_{j=1}^{j^*} \int \mathrm{d}M_{C,j}(u; H_0, \boldsymbol{\nu}) a_j\{u, \bar{\mathbf{V}}(u)\} I(T > u, Y \geq u). \quad (3.58)$$

This equation is used to construct estimating equations as it depends only on observables and is satisfied at the true values of $\boldsymbol{\nu}$ and H_0.

3.7.3 Parameter estimation under model B_q

In this section, the estimating equations used to find estimates of H_0, $\boldsymbol{\nu}$ and finally $S_T(t)^*$ are given. Again all these equations come from Rotnitzky et al. (2007).

Under model B_q, $\boldsymbol{\nu}$ has to be estimated before an estimator of $S_T(t^*)$ can be found. If H_0 were known, then $\boldsymbol{\nu}$ could be estimated using

$$\sum_i m(O_i; \boldsymbol{\nu}, H_0, d) = 0,$$

where $d_j\{u, \bar{\mathbf{V}}(u)\}$ are user-specified functions. Rotnitzky et al. (2007) say that a natural choice is

$$d_j\{u, \bar{\mathbf{V}}(u)\} = (\mathbf{0}', w_j(u, \bar{\mathbf{V}}(u))', \mathbf{0}')',$$

where the first and last $\mathbf{0}$ are zero vectors with $\sum_{l=1}^{j-1} \dim(\boldsymbol{\nu}_l)$ and $\sum_{l=j+1}^{j^*} \dim(\boldsymbol{\nu}_l)$ rows respectively. This form for $d_j\{u, \bar{\mathbf{V}}(u)\}$ is also used here.

As H_0 is unknown, then it has to be estimated. Equation 3.56 cannot be used to construct an estimator for H_0 as the RHS of the equation still depends on $\boldsymbol{\nu}$. However,

a profile estimator $\tilde{H}^{\boldsymbol{\nu}}_{0,j}(u)$ can be computed by solving the empirical version of (3.56) for each fixed $\boldsymbol{\nu}$, using that (3.56) is equivalent to

$$H^{\boldsymbol{\nu}}_{0,j}(u) = \int_0^u \frac{-\mathrm{d}F_{X|\Delta=0,J=j}(s)Pr(\Delta=0,J=j)}{E\left[\Delta I(T\geq s)\frac{\pi\{s|\bar{\mathbf{V}}(s),T;\boldsymbol{\nu},H^{\boldsymbol{\nu}}_0\}}{\pi\{T|\bar{\mathbf{V}}(T),T;\boldsymbol{\nu},H^{\boldsymbol{\nu}}_0\}}r_j\{s,\bar{\mathbf{V}}(s),T;\boldsymbol{\nu}_j\}\right]}.$$

So, $\tilde{h}^{\boldsymbol{\nu}}_{0,j}(u)=0$ if u is not a censoring time, and at each C_m, $\hat{h}_{0,j}$ is the solution to

$$\tilde{h}^{\boldsymbol{\nu}}_{0,j}(C_m) = \\ n^{(j)}_m\left[\sum_{i=1}^n \Delta_i I(T_i>C_m)\frac{\pi\{C_m^-|\bar{\mathbf{V}}_i(C_m^-),T_i;\boldsymbol{\nu},\tilde{H}^{\boldsymbol{\nu}}_0\}}{\pi\{T_i|\bar{\mathbf{V}}_i(T_i),T_i;\boldsymbol{\nu},\tilde{H}^{\boldsymbol{\nu}}_0\}}r_j\{C_m,\bar{\mathbf{V}}_i(C_m),T_i;\boldsymbol{\nu}_j\}\right]^{-1}$$

where, for each $\boldsymbol{\nu}$,

$$\pi\{u|\bar{\mathbf{V}}(u),T;\boldsymbol{\nu},\tilde{H}^{\boldsymbol{\nu}}_0\} = \prod_{j=1}^{j^*}\prod_{C_\ell|0<C_\ell\leq u}\left[1-\tilde{h}^{\boldsymbol{\nu}}_{0,j}(C_\ell)r_j\{C_\ell,\bar{\mathbf{V}}(C_\ell),T;\boldsymbol{\nu}_j\}\right],$$

and $\pi\{u^-|\bar{\mathbf{V}}(u^-),T;\boldsymbol{\nu},\tilde{H}^{\boldsymbol{\nu}}_0\}$ is defined as $\pi\{u|\bar{\mathbf{V}}(u),T;\boldsymbol{\nu},\tilde{H}^{\boldsymbol{\nu}}_0\}$ but with the second product ranging over all C_ℓ strictly less than u. The estimator $\tilde{h}^{\boldsymbol{\nu}}_{0,j}(C_m)$ needs to be computed recursively.

Now that we have an estimator $\tilde{H}^{\boldsymbol{\nu}}_0$, we can obtain an estimator $\hat{\boldsymbol{\nu}}$ of $\boldsymbol{\nu}$ using

$$\sum_{i=1}^n m(O_i;\boldsymbol{\nu},\tilde{H}^{\boldsymbol{\nu}}_0,d)=0.$$

Finally we can compute $\hat{S}_T(t^*)$ using

$$\sum_{i=1}^n \frac{\Delta_i}{\pi\{T_i|\bar{\mathbf{V}}_i(T_i),T_i;\hat{\boldsymbol{\nu}},\tilde{H}^{\hat{\boldsymbol{\nu}}}_0\}}\{I(T_i>t^*)-S_T(t^*)\}=0,$$

which is based on (3.55).

It is also possible to derive equations to give the variance but they are not presented here. The full derivation of equations for the variance is given in Rotnitzky et al. (2007).

3.7.4 Application of sensitivity analysis to the Liver Registration data set

In the Liver Registration dataset, we want to estimate $S_T(t^*) = Pr(T > t^*)$ for any $t^* \in (0,\kappa)$, where $T=\min(T^*,\kappa)$, $\kappa = 1260$ and T^* is the time from registration on the waiting list until death. So, $\kappa = 1260$ is used as the maximum follow up time observed was 1265 days as we do not want to discard too much data just to satisfy a technical issue. The choice of $\kappa = 1260$ is fairly arbitrary and we could have easily used $\kappa = 1261, 1262, 1263,$ or 1264 instead.

In our dataset, there are two competing censoring mechanisms, so $j^* = 2$. The first censoring time, $j = 1$, is the time to administrative censoring, which is assumed to be non-informative so q_1 is set to zero. The second censoring time $j = 2$, is the time to transplant and

$$q_2\{t, \bar{\mathbf{V}}(t), T; \omega, \zeta\} = \omega\{I(T < \kappa)(T - t) + I(T \geq \kappa)(\zeta - t)\}$$

is the censoring bias function for this cause of censoring. It is assumed that ω is known and takes values in the set $\{-1, -0.7, -0.5, -0.3, 0, 0.3, 0.5, 0.7, 1\}$. These values are chosen as ω is the assumed amount of residual dependence and we need a reasonable number of values ranging from 1 to -1 to assess the sensitivity of the estimator to different amounts of residual dependence. We set $\zeta = 1335$ days and this represents roughly the expected time until the event for subjects who have not experienced the event by 1260 days.

Negative values of ω are equivalent to assuming that, among patients who are at risk at time t and with the same covariate history up to t, those who would experience the the event earlier are more likely to be censored at time t than those who would experience the event later. Also when $\omega < 0$, the term $I(T \geq t)(\zeta - t)$ is equivalent to assuming that, among subjects who at risk at time t and with the same covariate history up to time t, the hazard of censoring at time t for whose who would experience the event after time κ is $\exp\{\omega(\zeta - \kappa)\}$ times smaller than the hazard of censoring for those who would experience the event just before time κ.

The covariates that are included in $\bar{\mathbf{V}}(u)$ are time-dependent UKELD and age. There are only two UKELD observations for patients who receive a transplant, these are taken at time of registration and time of transplant. Therefore linear interpolation is used to obtain values of UKELD at time values between these two points. For all other patients, there is only the UKELD score at time of registration, so this value is used at all times.

These two covariates are the only ones included in the model as the other possible covariates that could have been included were ethnicity and primary liver category, both of which are categorical variables with several levels. The dimensionality that these covariates add to the model would have made estimation of μ under the semi-parametric model much more computationally intensive.

The programs used to carry out the sensitivity analysis given in Rotnitzky et al. (2007) are available online from

`http://www.blackwellpublishing.com/rss.`

The programs available from this website are used here and amended slightly so that the sensitivity analysis could be carried out on the Liver Registration dataset.

The plot in Figure 3.12 shows the results of the sensitivity analysis. The thicker solid line is $\hat{S}_T(t^*)$ when it is assumed that $q_2 = 0$. This is the model that assumes sequential

ignorability of censoring after adjusting for time-dependent UKELD and recipient age. The Kaplan-Meier estimate of the survival function is included for comparison, this is given by the thinner solid line. The dotted lines on the plot are the estimates of the survival function when there is some negative residual dependence between time to event and time to censoring after adjusting for the covariates in $\bar{\mathbf{V}}(u)$. Similarly, the dashed lines are the estimates of the survival function when we assume the residual dependence is positive.

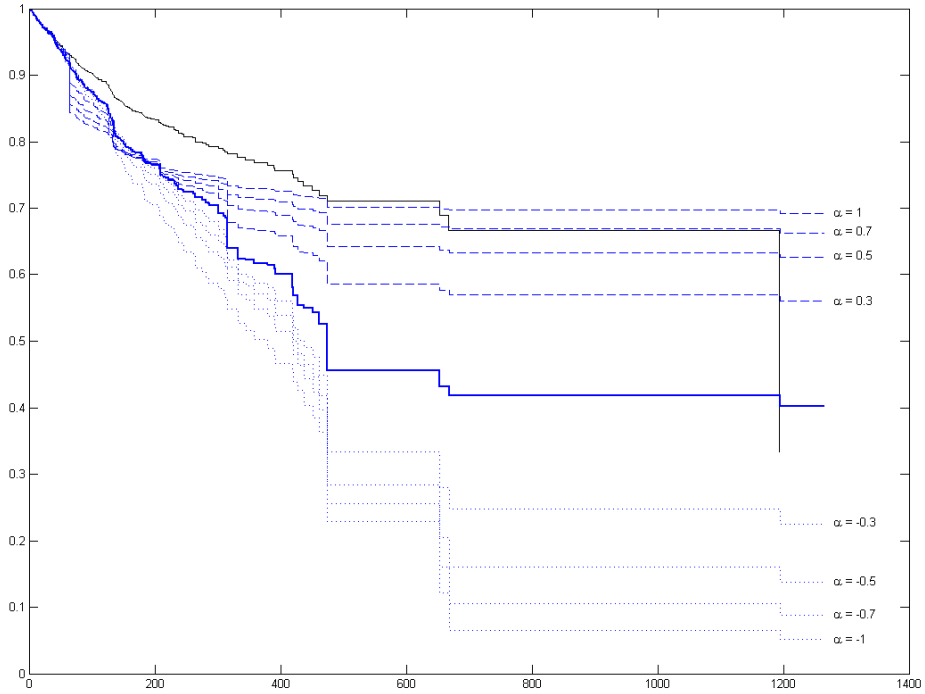

Figure 3.12: Plot showing estimated survival functions when fitting model B_q to the Liver Registration dataset for various values of α. The bolder solid line is the estimated survival function when fitting model B_q to the Liver Registration data set for $\alpha = 0$. The other solid line is the Kaplan-Meier estimate of the survival function.

The dashed and dotted lines on the plot show that the estimate of the survival function could change considerably if some of the values of the residual dependence that have been assumed are feasible. However, the covariates that have been included, particularly time-dependent UKELD, are significant predictors of time to death and time to censoring. So it is reasonable to assume that after adjusting for these covariates there can only be a small to moderate amount of residual dependence. Therefore it can reasonably be assumed that the residual dependence is likely to be in the interval $[-0.3, 0.3]$, which gives tighter

bounds on the estimate of the survival function. However, even these tighter bounds are too wide to be of much use in a practical application.

The estimator that assumes $q_2 = 0$, which is the assumption of sequential ignorability is not similar to the IPCW estimators fitted in Section 3.1.4, despite using a similar weighted approach. This is not just due to different covariates being used in the models for time to censoring for each of the estimators. It is because the estimator used here is not the KM estimator, instead it is

$$\hat{S}_T(t^*) = \frac{\sum_{i=1}^n I(T_i \leq C_i) I(T_i > t^*) / \pi\{T_i | \bar{\mathbf{V}}_i(T_i), T_i; \hat{\boldsymbol{\nu}}, \tilde{H}_0^{\hat{\boldsymbol{\nu}}}\}}{\sum_{i=1}^n I(T_i \leq C_i) / \pi\{T_i | \bar{\mathbf{V}}_i(T_i), T_i; \hat{\boldsymbol{\nu}}, \tilde{H}_0^{\hat{\boldsymbol{\nu}}}\}},$$

which is the number of failures yet to occur weighted by $\pi\{T_i | \bar{\mathbf{V}}_i(T_i), T_i; \hat{\boldsymbol{\nu}}, \tilde{H}_0^{\hat{\boldsymbol{\nu}}}\}$, over the total number of failures weighted by $\pi\{T_i | \bar{\mathbf{V}}_i(T_i), T_i; \hat{\boldsymbol{\nu}}, \tilde{H}_0^{\hat{\boldsymbol{\nu}}}\}$. In comparison, the IPCW KM estimator derived in Section 3.1.1 is

$$\hat{S}_T(t) = \prod_{\{i; t_i < t\}} \left(1 - \frac{\Delta_i \hat{W}_i(t_i)}{\left\{ \sum_{k=1}^n R_k(t_i) \hat{W}_k(t_i) \right\}} \right).$$

If the weight used is $\hat{W}_i(t) = 1/\hat{K}_i^{\mathbf{V}}(t)$ where

$$\hat{K}_i^{\mathbf{V}}(t) = \prod_{\{j; t_j < t, \Delta_j = 0\}} [1 - \hat{h}_0(t_j) \exp\{\hat{\boldsymbol{\beta}}_C' \mathbf{V}_i(t_j)\}],$$

then it appears that similar weights are used for the IPCW KM estimator and the estimator presented in this section when $\omega = 0$. However, the IPCW KM estimator uses Breslow's estimator of the hazard function, which is not used by Rotnitzky et al. (2007) when estimating the hazard function.

3.8 Summary

In Chapter 2, estimators that could be used to give bounds on the estimated survival function were reviewed. In this chapter, alternative approaches that can be used when there is potentially informative censoring in a data set are reviewed. These include approaches that use regression models for the censoring processes and sensitivity analyses for the parameters of non-informative censoring models. Generally, the methods presented in this chapter allow covariates to be incorporated much more easily than the estimators considered in Chapter 2. A summary of the advantages and disadvantages of the estimators that use models for the censoring process is given in Table 3.9. Also, the advantages and disadvantages of the sensitivity analyses reviewed in this chapter are summarised in Table 3.10. All the approaches given in this chapter share a disadvantage that is not included in Tables 3.9 and 3.10. This is that they all rely on some untestable assumption about the

nature of the dependence between T and C, due to the identifiability issues described in Section 1.1.2.

The most widely used approach that uses a regression model for the censoring process weights the contribution of individuals in the data set by the inverse of the probability of the individual remaining uncensored under this model. If these weighted contributions are used in the standard methods, such as the KM estimate of the survival function or Cox's proportional hazards model, then estimates of the survival function or parameters in the absence of any censoring can be obtained. The various models that are considered for the censoring process are Cox's proportional hazards model, Weibull proportional hazards model and Aalen's additive hazard model. We recommend use of a Cox proportional hazards model as it can be easily fitted using standard software and can also incorporate time-dependent covariates fairly easily. These inverse probability of censoring weighting methods use the assumption of sequential ignorability of censoring. This means that if all the prognostic factors for both T and C are adjusted for in the model for censoring, it can be assumed that C would then be independent of T. However, it is possible that some of the dependence between T and C is due to unmeasured factors, which is called residual dependence.

In Sections 3.3 to 3.6, methods are described that assess the sensitivity of parameter estimates from standard models to the assumption of informative censoring.

Firstly, two sensitivity analyses that use parametric survival models are considered. Both of these methods are computationally simple but cannot be used for a wide range of data sets as they use only standard parametric survival models. The change in parameter estimates for both of these sensitivity analyses can be expressed in the same form, which is the correlation coefficient of T and C multiplied by a sensitivity index, which allows direct comparison of the two methods. We recommend using the sensitivity analysis given in Siannis et al. (2005) and Siannis (2004) for parametric survival models as it allows estimation of the change in individual parameter estimates for covariates unlike the sensitivity analysis from Zhang and Heitjan (2006) which only allows use of a linear predictor. Also the sensitivity analysis in Zhang and Heitjan (2006) gives values of the sensitivity index that seem unfeasibly large.

Then, two sensitivity analyses for the Cox proportional hazards model are given. These approaches are much more computationally intensive than those for parametric models but can be applied to a greater number of data sets as the Cox proportional hazards model is more flexible than standard parametric survival models. The sensitivity analysis in Siannis (2011) is more computationally intensive than the sensitivity analysis in Siannis et al. (2005) and Siannis (2004) as it requires estimation of the baseline hazard function. However, the sensitivity analysis in Huang and Zhang (2008) is much more computationally

intensive than that in Siannis (2011). It also requires more untestable assumptions as a copula function needs to be specified as well as the level of dependence between T and C. It is for these reasons that we recommend use of the sensitivity analysis in Siannis (2011) for the Cox proportional hazards model.

Finally, a sensitivity analysis for an estimator that already accounts for informative censoring is considered. This is derived in Rotnitzky et al. (2007) and considers the sensitivity to residual dependence of an estimator that assumes sequential ignorability of censoring. Unfortunately, this estimator is not the same as the inverse probability of censoring weighted estimators considered previously. A semi-parametric model containing prognostic factors for T and C is used for C. Weights using the survival function from this semi-parametric model are used when deriving estimators of the marginal survival function for T. This approach can then be used to give bounds on the estimator that assumes sequential ignorability of censoring for different amounts of residual dependence. The drawbacks of this method are that the bounds derived are often too wide to be of use in a practical setting and it is so computationally intensive that it is not easy to include lots of covariates or factors with many levels.

The literature review carried out in this chapter has several important conclusions. The first is that the inverse probability of censoring weighted estimators presented in Section 3.1 are the most appropriate estimators for use in practical applications that we have found in the literature. Therefore, similar weights will be used when developing the survival benefit methodology in Chapter 6. Secondly, the sensitivity analyses described in Sections 3.3 to 3.6 are the most useful methodologies in the literature for assessing the sensitivity of results from standard models to the assumption of informative censoring. However, each of these methodologies have disadvantages that affect its usefulness is a practical setting. This is our motivation for the sensitivity analysis for piecewise parametric survival models derived in Chapter 4, which has the flexibility of the sensitivity analyses for Cox's proportional hazards model while retaining the computational simplicity of the sensitivity analyses for standard parametric survival models.

	Advantages	Disadvantages
IPCW estimators	• Uses sequential ignorability of censoring assumption which is fairly easy to understand and it seems intuitive that dependence would be caused by shared prognostic factors. • The Cox model for time to censoring can be fitted using standard software and it does not require much computation to obtain weights. • Most standard software can easily incorporate weights into Cox models.	• It is not possible to establish whether the correct model for censoring has been used. • There may be residual dependence that is not explained by shared prognostic factors, which would result in IPCW estimates being biased.
Weighted estimators using Aalen's additive hazard model	• Uses sequential ignorability of censoring assumption which is fairly easy to understand and it seems intuitive that dependence would be caused by shared prognostic factors. • Uses Aalen's additive hazard model which is more flexible than Cox's proportional hazards model.	• It is not possible to establish whether the correct model for censoring has been used. • There may be residual dependence that is not explained by shared prognostic factors, which would result in IPCW estimates being biased. • It is more difficult to fit Aalen's additive hazard model than Cox's proportional hazards model using standard software.

Table 3.9: Summary of the advantages and disadvantages of the estimators that use models of the censoring process that are reviewed in Chapter 3.

	Advantages	Disadvantages
Sensitivity analysis in Siannis et al. (2005) and Siannis (2004)	• Computation of sensitivity analysis equations is fairly simple. • Allows estimation of change in individual parameter estimates as well as a linear predictor.	• Only allows use of standard parametric survival models, which means it can only be applied to a restricted number of data sets.
Sensitivity analysis in Zhang and Heitjan (2006)	• Computation of sensitivity analysis equations is fairly simple.	• Only allows use of standard parametric survival models. • Can only be carried out on scalar parameters or linear predictors. • Gives values of the sensitivity index that seem unfeasibly large.
Sensitivity analysis in Huang and Zhang (2008)	• Uses Cox's proportional hazards model, which is more flexible than standard parametric survival models. • Allows estimation of the change in individual parameter estimates.	• Very computationally intensive, much more than the sensitivity analysis in Siannis (2011). • Requires more untestable assumptions than other methods reviewed in Chapter 3, as a copula needs to be chosen as well as the level of dependence.
Sensitivity analysis in Siannis (2011)	• Uses Cox's proportional hazards model, which is more flexible than standard parametric survival models. • Allows estimation of the change in individual parameter estimates.	• More computationally intensive than the sensitivity analyses in Siannis et al. (2005), Siannis (2004) and Zhang and Heitjan as it requires estimation of the baseline hazard function.
Sensitivity analysis in Rotnitzky et al. (2007)	• Allows assessment of the sensitivity to residual dependence of an estimator that assumes sequential ignorability of censoring.	• The estimator used is not the same as the IPCW estimator, which is more widely used. • Very computationally intensive, it is not possible to include lots of covariates or factors with many levels. • The bounds on the estimate of the survival function found are often too wide to be of use practically.

Table 3.10: Summary of the advantages and disadvantages of the sensitivity analyses reviewed in Chapter 3.

Chapter 4

Sensitivity Analysis for Informative Censoring in Piecewise Exponential Models

Sensitivity analyses that estimate how the results from fitting standard models would change in the presence of informative censoring are useful, due to the identifiability issues that we face. It is for this reason that here we present a sensitivity analysis method that is not only suited to our particular setting, but could also be applied to many other situations.

The method allows us to estimate the change in the parameter estimates for a piecewise exponential model when we assume a small amount of informative censoring instead of non-informative censoring. This extends the sensitivity analysis in Siannis et al. (2005) and Siannis (2004), which only considered standard parametric models. The method they present is appealing as it is easy to apply, but it could be improved by extending the range of models to which it applies. We chose to work with piecewise exponential models as by using sensible cutpoints to split the study time into intervals and assuming constant hazards in each interval, we can approximate a wide range of baseline hazard functions.

A sensitivity analysis that uses the same assumption for the association between T and C is given in Siannis (2011). However, this sensitivity analysis is for the Cox proportional hazards model instead of a parametric model. The sensitivity analysis presented here retains the computational simplicity of the parametric analyses of Siannis (2004) and Siannis et al. (2005) whilst enjoying the flexibility of the approach of Siannis (2011).

We will first outline the sensitivity analysis for a piecewise exponential model with a scalar parameter in each interval. The sensitivity analysis will then be extended so that covariates can be included.

4.1 Notation and Model Specification

We are interested in the joint distribution of T, the time to failure variable and C, the time to censoring variable, so we can assess the dependence between the two. However, we only observe $Y = \min(T, C)$ along with an indicator function $I = 1$ if $T \leq C$ and $I = 0$ otherwise. This means that we must make additional assumptions before we can identify the joint distribution.

A piecewise exponential model will be used for the marginal distributions of both T and C. We split the study time into intervals and assume a constant hazard in each interval. This approach was introduced in Breslow (1974); it is what is now called the piecewise exponential model. This should give us greater flexibility than the standard exponential and Weibull models, as we can approximate most hazard functions.

As we have introduced intervals into the model, we use a piecewise approach to obtain the log-likelihood. We only have the observation time, y_i, for each individual and the piecewise approach requires a time variable corresponding to each interval for each individual. Therefore we define the exposure time for individual i in interval j, which is

$$\begin{aligned} y_{ij} &= a_j - a_{j-1} \qquad && j = 1, \ldots, N_i - 1 \\ y_{ij} &= y_i - a_{N_i - 1} \qquad && j = N_i, \end{aligned}$$

where a_j is the upper endpoint of the jth interval. The lower endpoint of the first interval is $a_0 = 0$. Here N_i denotes the number of the interval in which individual i experiences either failure or the censoring of interest at time y_i. Once having experienced one of these events, individual i has no further exposure in later intervals.

Therefore, there are now three possible times that may be observed for each individual at risk in any of the intervals. These are T, the failure time, C_I, the censoring that occurs within an interval, and C_E, the censoring at the end of an interval. We will treat the censoring at the end of each interval, which has been introduced by the use of a piecewise model, as independent of any censoring that takes place in the intervals. This censoring is similar to end of study censoring, which is also usually treated as non-informative censoring.

Two indicator variables are needed, first to distinguish between a failure time and a censored time and then to distinguish between the two different types of censoring. These indicator variables are

$$I_{ij} = \begin{cases} 1, & \text{if } i\text{th individual fails in } j\text{th interval} \\ 0, & \text{if } i\text{th individual does not fail in } j\text{th interval} \end{cases}$$

and, when $I_{ij} = 0$,

$$Z_{ij} = \begin{cases} 1, & \text{when individual } i \text{ censored before the end of interval } j \\ 0, & \text{when individual } i \text{ censored at the end of interval } j. \end{cases}$$

As we are using a piecewise exponential model, we can take advantage of the lack of memory between the intervals. If we condition on $T > a_{j-1}$, then the survivor function $S_T(t|T > a_{j-1}) = S_T(t - a_{j-1})$ for the jth interval depends only on the parameter associated with that interval, θ_j, and the mean of the distribution in the jth interval is given by θ_j^{-1}. Let $t_j = t - a_{j-1}$ be the amount of time passed in the jth interval, then the survivor function can be denoted by $S_T(t_j, \theta_j)$. For the ith individual, the survivor function for the jth interval would be $S_T(y_{ij}, \theta_j)$. The density, hazard and integrated hazard functions for T in the jth interval,

$$f_T(t_j, \theta_j) = -\frac{d}{dt} S_T(t_j, \theta_j), \quad h_T(t_j, \theta_j) = -\frac{d}{dt} \log S_T(t_j, \theta_j),$$
$$\text{and} \quad H_T(t_j, \theta_j) = -\log S_T(t_j, \theta_j),$$

also only depend on θ_j. The score and information functions for the density function $f_T(t_j, \theta_j)$ are defined by

$$s_T(t, \theta_j) = \frac{\partial}{\partial \theta_j} \log f_T(t_j, \theta_j) \quad \text{and} \quad i_{\theta_j} = \text{Var}_T\{s_T(T, \theta_j)\}.$$

Similarly, if we condition on $C > a_{j-1}$ then the survivor function $S_C(c|C > a_{j-1})$ for C_I in the jth interval only depends on the nuisance parameter, γ_j. For ease of notation, without ambiguity subscript C will be used for functions relating to C_I. Let $c_j = c - a_{j-1}$ and the survivor function for C_I in the jth interval can be denoted $S_C(c_j, \gamma_j)$. There are the corresponding functions $f_C(c_j, \gamma_j)$, $h_C(c_j, \gamma_j)$, $H_C(c_j, \gamma_j)$, $s_C(c, \gamma_j)$ and i_{γ_j} for C_I.

It is now necessary to make an assumption concerning the conditional distribution of C_I given T, so that we can identify the joint distribution of T and C_I. As in Siannis et al. (2005), Siannis (2004) and Siannis (2011), we assume that the conditional distribution of C_I given T has the same parametric distribution as the marginal distribution of C_I. However, the parameter of the conditional density is allowed to depend on T. Therefore, the conditional density in the jth interval can be written explicitly as

$$f_{C|T}(c_j|t, \gamma_j, \delta, \theta_j) = f_C(c_j, \gamma_j + \delta i_{\gamma_j}^{-1/2} B(t_j, \theta_j)),$$

where i_{γ_j} is the information function for C_I. The dependence between T and C_I is defined by δ and $B(t_j, \theta_j)$. These can be thought of as a correlation coefficient, that quantifies the amount of dependence between the two processes, and a bias function which gives a form to this dependence. More specifically, $B(t_j, \theta_j)$ quantifies the dependence between T

and censoring just after time t, for those who remain at risk at time t, as discussed in 3.3. The choice of the form of the bias function that we will use in this method is discussed in Section 4.2.1.

As we will let the parameters vary between the intervals, we will have the vectors $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}$ with θ_j and γ_j being the scalar parameters in the jth of the m intervals in our model.

4.2 Development of Sensitivity Analysis

Here we describe the development of a sensitivity analysis that can be applied to piecewise data. This is an extension of the approaches set out in Siannis et al. (2005) and Siannis (2004). At first it will not incorporate covariates but it will be shown that it can be extended to do so in Section 4.2.3.

Let $\ell_\delta(\boldsymbol{\theta}, \boldsymbol{\gamma})$, be the log-likelihood function when T and C_I are dependent as outlined above in Section 4.1. Then

$$\ell_\delta(\boldsymbol{\theta}, \boldsymbol{\gamma}) = \sum_{i=1}^{n}\sum_{j=1}^{m}\Big\{I_{ij}\log K_1(y_{ij}) + Z_{ij}(1-I_{ij})\log K_2(y_{ij}) + (1-I_{ij})(1-Z_{ij})\log K_3(y_{ij})\Big\}, \tag{4.1}$$

where

$$\begin{aligned} K_1(y_{ij}) &= \int_{y_{ij}}^{\infty} f_{T,C}(y_{ij}, u)du \\ K_2(y_{ij}) &= \int_{y_{ij}}^{\infty} f_{T,C}(u, y_{ij})du \\ \text{and} \quad K_3(y_{ij}) &= \int_{y_{ij}}^{\infty}\int_{y_{ij}}^{\infty} f_{T,C}(t, c)dt\, dc. \end{aligned} \tag{4.2}$$

These can be thought of as the likelihood contributions for each of the three types of observations that may occur in each interval. To avoid having integrals in the above contributions that cannot be evaluated analytically, the joint density for T and C_I in the jth interval is written

$$\begin{aligned} f_{T,C}(t_j, c_j) &= f_T(t_j, \theta_j) f_C(c_j, \gamma_j + \delta i_{\gamma_j}^{-1/2} B(t_j, \theta_j)) \\ &\simeq f_T(t_j, \theta_j) f_C(c_j, \gamma_j)[1 + \delta i_{\gamma_j}^{-1/2} s_C(c_j, \gamma_j) B(t_j, \theta_j)]. \end{aligned} \tag{4.3}$$

Now that the model has been fully specified, it is possible to find approximations of the contributions in (4.2). Once these have been substituted into (4.1), the log likelihood

becomes:

$$\ell_\delta(\boldsymbol{\theta},\boldsymbol{\gamma}) \simeq \ell_0(\boldsymbol{\theta},\boldsymbol{\gamma}) - \delta \sum_{i=1}^{n}\sum_{j=1}^{m} i_{\gamma_j}^{-1/2}\Big\{ I_{ij}B(y_{ij},\theta_j)\frac{\partial}{\partial\gamma_j}H_C(y_{ij},\gamma_j) + (1-I_{ij})(1-Z_{ij})\frac{\partial}{\partial\gamma_j}H_C(y_{ij},\gamma_j)\mu(y_{ij},\theta_j) - Z_{ij}(1-I_{ij})s_C(y_{ij},\gamma_j)\mu(y_{ij},\theta_j)\Big\}, \tag{4.4}$$

where

$$\mu(y_{ij},\theta_j) = \frac{\displaystyle\int_{y_{ij}}^{\infty} B(u,\theta_j)f_T(u,\theta_j)du}{S_T(y_{ij},\theta_j)}$$

and

$$\ell_0(\boldsymbol{\theta},\boldsymbol{\gamma}) = \sum_{i=1}^{n}\sum_{j=1}^{m}\Big\{ I_{ij}\log h_T(y_{ij},\theta_j) + Z_{ij}(1-I_{ij})\log h_C(y_{ij},\gamma_j) - H_T(y_{ij},\theta_j) - H_C(y_{ij},\gamma_j)\Big\}. \tag{4.5}$$

For a fixed value of δ, $\hat{\boldsymbol{\theta}}_\delta$ is the vector of values that maximises (4.4). Note that the first term in (4.4), $\ell_0(\boldsymbol{\theta},\boldsymbol{\gamma})$, is the log-likelihood in the non-informative censoring model.

To be able to assess how much the parameter estimates change under the assumption of dependent censoring, an estimate of the difference between them is needed. The estimate of θ_j under the assumption of dependent censoring in the jth interval is denoted by $\hat{\theta}_{\delta j}$. Similarly the estimated value of the parameter under independent censoring in the jth interval is denoted as $\hat{\theta}_{0j}$. To be able to obtain an approximation of the difference between these two values, it is necessary to use Taylor expansions about θ_j of the score functions

$$r_0(\hat{\theta}_{0j}) = \frac{\partial}{\partial\theta_j}\ell_0(\boldsymbol{\theta},\boldsymbol{\gamma})\Big|_{\hat{\theta}_{0j}} \quad \text{and} \quad r_\delta(\hat{\theta}_{\delta j}) = \frac{\partial}{\partial\theta_j}\ell_\delta(\boldsymbol{\theta},\boldsymbol{\gamma})\Big|_{\hat{\theta}_{\delta j}}. \tag{4.6}$$

These are the score functions for the jth interval under the assumption of independent and dependent censoring respectively. Therefore they are the score functions for the likelihoods given in (4.4) and (4.5) respectively.

The score functions given in (4.6) are expanded about θ_j and set equal to zero to give

$$\begin{aligned} r_0(\hat{\theta}_{0j}) &\simeq r_0(\theta_j) - (\hat{\theta}_{0j}-\theta_j)i_j(\boldsymbol{\theta}) = 0 \\ r_\delta(\hat{\theta}_{\delta j}) &\simeq r_\delta(\theta_j) - (\hat{\theta}_{\delta j}-\theta_j)i_j(\boldsymbol{\theta}) = 0 \end{aligned} \tag{4.7}$$

where

$$i_j(\boldsymbol{\theta}) = -\frac{\partial^2}{\partial\theta_j^2}\ell_0(\boldsymbol{\theta},\boldsymbol{\gamma}).$$

Rearranging the two equations in (4.7) gives

$$(\hat{\theta}_{\delta j}-\hat{\theta}_{0j})i_j(\boldsymbol{\theta}) \simeq r_\delta(\theta_j) - r_0(\theta_j).$$

So, an approximation of the difference between the parameter estimates is given by

$$\begin{aligned}\hat{\theta}_{\delta j} - \hat{\theta}_{0j} \simeq \delta(i_j(\boldsymbol{\theta}))^{-1} \sum_{i \in R_j} i_{\gamma_j}^{-1/2} \Big\{ & Z_{ij}(1 - I_{ij}) s_C(y_{ij}, \gamma_j) \frac{\partial \mu(y_{ij}, \theta_j)}{\partial \theta_j} \\ & - (1 - Z_{ij})(1 - I_{ij}) \frac{\partial H_C(y_{ij}, \gamma_j)}{\partial \gamma_j} \frac{\partial \mu(y_{ij}, \theta_j)}{\partial \theta_j} \\ & - I_{ij} \frac{\partial H_C(y_{ij}, \gamma_j)}{\partial \gamma_j} \frac{\partial B(y_{ij}, \theta_j)}{\partial \theta_j} \Big\} \end{aligned} \tag{4.8}$$

for the jth interval, where R_j is the set of individuals who are at risk in the jth interval. We see that in (4.8) there are parameter estimates on the LHS and parameters on the RHS of the expression. This is a result of rearranging the two equations in (4.7). It means that when the sensitivity analysis is applied, sensible estimates of the parameters must be substituted into the RHS of (4.8).

4.2.1 Choice of $B(t, \theta)$

The argument presented in this section gives a general method for choosing the bias function $B(t_j, \theta_j)$ in the jth interval. This is adapted from the argument that was presented in Siannis et al. (2005) as justification for the choice of bias function. For simplicity, we will look at the case where there are just scalar parameters, θ_j and γ_j in each interval. We shall assume that non-ignorability comes from the correlation between individual-specific random effects in the distributions of T and C. Then for a given patient, T and C would be independent given the random effects with density functions given by

$$g_T(t_j, \theta_j + \epsilon_T i_{\theta_j}^{-1/2}) \quad \text{and} \quad g_C(c, \gamma_j + \epsilon_C i_{\gamma_j}^{-1/2}),$$

where ϵ_T and ϵ_C are random effects with mean zero, variances σ_T^2 and σ_C^2 and covariance σ_{TC}. We shall assume that all three of these second moments are fairly small, with the same order of magnitude. This will allow the use of Taylor expansions around $\epsilon_T = 0$ and $\epsilon_C = 0$, where we ignore terms that are above second order. These can be used to gain approximations to the marginal distributions of T and C where

$$\begin{aligned} f_T(t_j, \theta_j) &= E\left[g_T(\theta_j + \epsilon_T i_{\theta_j}^{-1/2}) \right] \\ &\simeq g_T(t_j, \theta_j) + \frac{\sigma_T^2}{2 i_{\theta_j}} \frac{\partial^2 g_T(t_j, \theta_j)}{\partial \theta_j^2} \end{aligned}$$

and similarly

$$f_C(c_j, \gamma_j) \simeq g_C(c_j, \gamma_j) + \frac{\sigma_C^2}{2 i_{\gamma_j}} \frac{\partial^2 g_C(c_j, \gamma_j)}{\partial \gamma_j^2}.$$

Also, an approximation of the joint distribution can be found from

$$f_{T,C}(t_j, c_j) = E\left[g_T(t_j, \theta_j + \epsilon_T i_{\theta_j}^{-1/2}) g_C(c, \gamma_j + \epsilon_C i_{\gamma_j}^{-1/2}) \right].$$

Once the Taylor expansions have been multiplied out and we have used the fact that we can write

$$f_T(t_j,\theta_j)f_C(c_j,\gamma_j) \simeq g_T(t_j,\theta_j)g_C(c_j,\gamma_j) + \frac{\sigma_C^2}{2i_{\gamma_j}} g_T(t_j,\theta_j)\frac{\partial^2}{\partial\gamma_j^2} g_C(c_j,\gamma_j) + \frac{\sigma_T^2}{2i_{\theta_j}} g_C(c_j,\gamma_j)\frac{\partial^2}{\partial\theta_j^2} g_T(t_j,\theta_j)$$

then the joint density can be written as

$$f_{T,C}(t_j,c_j) \simeq f_T(t_j,\theta_j)f_C(c_j,\gamma_j)[1+\sigma_{TC}(i_{\theta_j}i_{\gamma_j})^{-1/2}s_T(t_j,\theta_j)s_C(c_j,\gamma_j)]. \tag{4.9}$$

If we compare (4.9) to (4.3), then we can see that with δ appropriately defined, the two equations will be equal if

$$B(t_j,\theta_j) = i_{\theta_j}^{-1/2} s_T(t_j,\theta_j).$$

Other justifications for choosing this form of $B(t_j,\theta_j)$ may be given. The two justifications given here are from Siannis (2011). Firstly, the form of the dependence here is completely unknown so any assumptions made about it should be as weak as possible as far as information about θ_j is concerned. There is also a nice symmetry in the competing risks set-up if this $B(t_j,\theta_j)$ is used. It means the conditional distribution of C given T has the same form as the conditional distribution of T given C.

4.2.2 Proportional hazards structure

As in Siannis et al. (2005) and Siannis (2004), we use a proportional hazards structure to simplify our model so that the hazard functions of T and C_I have the form

$$h_T(t_j,\theta_j) = e^{\theta_j}h_T^*(t_j) \quad \text{and} \quad h_C(c_j,\gamma_j) = e^{\gamma_j}h_C^*(c_j),$$

where $h_T^*(t_j)$ and $h_C^*(c_j)$ are baseline hazard functions. Consequently,

$$s_T(t_j,\theta_j) = 1 - H_T(t_j,\theta_j), \quad s_C(c_j,\gamma_j) = 1 - H_C(c_j,\gamma_j) \quad \text{and} \quad i_{\theta_j} = i_{\gamma_j} = 1. \tag{4.10}$$

If we take $B(t_j,\theta_j)$ to be the standardized score function, the reasoning for which was outlined in Section 4.2.1,

$$B(t_j,\theta_j) = i_{\theta_j}^{-1/2} s_T(t_j,\theta_j) \tag{4.11}$$

then we can combine (4.10) and (4.11) to give

$$B(t_j,\theta_j) = 1 - H_T(t_j,\theta_j) \qquad \text{and} \qquad \mu(t_j,\theta_j) = H_T(t_j,\theta_j). \tag{4.12}$$

This proportional hazards structure also allows us to give simple expressions for the partial derivatives in (4.8) as

$$\frac{\partial}{\partial\theta_j}H_T(t_j,\theta_j) = H_T(t_j,\theta_j) \quad \text{and} \quad \frac{\partial}{\partial\gamma_j}H_C(c_j,\gamma_j) = H_C(c_j,\gamma_j). \tag{4.13}$$

If we now apply the proportional hazards structure to (4.8), it will simplify greatly to give

$$\hat{\theta}_{\delta j} - \hat{\theta}_{0j} \simeq \delta i_j(\boldsymbol{\theta})^{-1} \sum_{i \in R_j} \left\{ H_T(y_{ij}, \theta_j) H_C(y_{ij}, \gamma_j) - Z_{ij}(1 - I_{ij}) H_T(y_{ij}, \theta_j) \right\}, \qquad (4.14)$$

which applies to the jth interval, where

$$i_j(\boldsymbol{\theta}) = -\frac{\partial^2}{\partial \theta_j^2} \ell_0(\boldsymbol{\theta}, \boldsymbol{\gamma}) = \sum_{i \in R_j} H_T(y_{ij}, \theta_j). \qquad (4.15)$$

Notice that it is necessary to perform the sensitivity analysis separately on the parameters for each interval. We are justified to use the proportional hazards structure as long as the proportional hazards assumption holds within each interval. The piecewise exponential model satisfies this as the hazard functions for T and C_I are piecewise constant. These piecewise constant hazards can provide a fair approximation to most hazard functions provided sensible cut-points for the intervals are identified. Large intervals may be used when the hazard function is changing slowly. When it is changing rapidly small intervals would capture this better. This gives more flexibility than the Weibull model assumed in Siannis (2004) because the hazard for a Weibull distribution has to be monotonic, and there is no such restriction when using a piecewise constant hazard.

4.2.3 Inclusion of covariates

Siannis et al. (2005) also show how covariates can be included in the sensitivity analysis approach. This has been briefly discussed previously in Section 3.3. A similar approach is used in this section to incorporate covariates into the sensitivity analysis for piecewise parametric models. Siannis et al. (2005) derive an equation that approximates the value of $\hat{\boldsymbol{\theta}}_\delta - \hat{\boldsymbol{\theta}}_0$, where $\boldsymbol{\theta}$ is the vector of parameters for the covariate vector $\mathbf{x}$ that replaces the scalar parameter θ. However, when applying their method to data they only consider the change in the linear predictor $w(\mathbf{x}) = \boldsymbol{\theta}_j' \mathbf{x}$ as it is computationally simpler. In this section, we will consider both approaches and derive equations for a sensitivity analysis for $\boldsymbol{\theta}$ and a sensitivity analysis for $w(\mathbf{x})$.

In order to incorporate covariates in the sensitivity analysis for piecewise parametric models, we replace the scalar parameters θ_j and γ_j in the jth interval by $\boldsymbol{\theta}_j' \mathbf{x}$ and $\boldsymbol{\gamma}_j' \mathbf{x}$. However, $\boldsymbol{\gamma}$ is a nuisance parameter so we will introduce the scalar $\eta_j = \boldsymbol{\gamma}_j' \mathbf{x}$. It is due to its dependence on $\boldsymbol{\gamma}_j$, that η is also dependent on j. The use of piecewise exponential models with covariates is described in Friedman (1982). The hazard function of the ith individual in the jth interval is defined to be

$$h_{ij} = \exp(\alpha_j + \sum_{k=1}^{p} \beta_k x_{ik}). \qquad (4.16)$$

This is the same as splitting $\boldsymbol{\theta}_j'\mathbf{x}$ into an intercept for each interval, α_j, and a component for the p covariates included in the model, given by $\sum_{k=1}^{p}\beta_k x_{ik}$, which remains constant over the intervals.

Firstly, we derive the sensitivity analysis for the function $w_j(\mathbf{x}) = \boldsymbol{\theta}_j'\mathbf{x}$ rather than $\boldsymbol{\theta}_j$. This is done using a method similar to that used in Section 4.2 but with expansions of the score functions

$$\boldsymbol{r}_0(w_{0j}(\mathbf{x})) = \frac{\partial}{\partial w_{0j}(\mathbf{x})}\ell_0(w_j(\mathbf{x}), z_j(\mathbf{x})), \quad \text{and}$$
$$\boldsymbol{r}_\delta(w_{\delta j}(\mathbf{x})) = \frac{\partial}{\partial w_{\delta j}(\mathbf{x})}\ell_\delta(w_j(\mathbf{x}), z_j(\mathbf{x})).$$

Note that the function $z(\mathbf{x}) = \eta$ does not contain the parameter of interest. Then the difference in the linear predictors would be given by:

$$\hat{w}_\delta(\mathbf{x}) - \hat{w}_0(\mathbf{x}) \simeq \delta\frac{\sum_{i\in R_j} H_T(y_{ij}, \boldsymbol{\theta}, \mathbf{x})[H_C(y_{ij}, \eta_{ij}) - (1 - I_{ij})Z_{ij}]}{\sum_{i\in R_j} H_T(y_{ij}, \boldsymbol{\theta}, \mathbf{x})}, \tag{4.17}$$

using $i(\boldsymbol{\theta}, \tilde{\mathbf{x}}) = \sum_{i=1}^{n}\sum_{j=1}^{m} H_T(y_{ij}, \boldsymbol{\theta}, \mathbf{x})$ and where R_j is the risk set in the jth interval.

Deriving a sensitivity analysis for $\boldsymbol{\theta}$ is much simpler if the model is expressed as a vector of parameters that stays constant over all the intervals. This is possible for the piecewise exponential model and we will now show how to reformulate the parameter vector.

When including covariates there is a vector of parameters $\boldsymbol{\theta}_j$ for each interval. Only the intercept component changes across intervals. If we consider the case without other covariates, then this means we can express the log-hazard of failure for the ith individual in the jth interval as

$$\log h_T(t) = \theta_j$$

It is possible to fit the same model with a vector of parameters $\boldsymbol{\theta}$, that remains constant over all the intervals. This is achieved by specifying a constant intercept over all intervals, θ_0, along with a factor, $\mathbf{v}_j$, that indicates the interval under consideration. The log-hazard defined above can then be expressed as

$$\log h_T(t) = \theta_0 + \boldsymbol{\theta}'\mathbf{v}_j$$

The parameter estimates of this factor correspond to the contrasts between the intercept in a given interval and the baseline intercept. As this approach means that we will always have a vector of parameters, it is trivial to consider standard covariates as well. However, we now need to make clear the dependence of the vector of covariates on both i and j. So to remain consistent with previous notation we define $\mathbf{x}_{ij}$ to be the vector of covariates for individual i in interval j.

We define the model for the censoring time variable, C_I, in the same way. We will continue to consider the scalar $\eta_{ij} = \boldsymbol{\gamma}'\mathbf{x}_{ij}$ for the ith individual in the jth interval, as $\boldsymbol{\gamma}$ is a nuisance parameter.

The lack of memory between intervals is still being used so the functions for T and C_I are still conditioned on the individual surviving beyond time a_{j-1}.

In a sample of n observations with m intervals, $\tilde{\mathbf{x}}$ is the array containing the $\mathbf{x}_{ij}$ vectors and $\boldsymbol{\eta}$ is the matrix with the η_{ij} as its elements. So the log-likelihood when we include covariates becomes:

$$\begin{aligned}\ell_\delta(\boldsymbol{\theta}, \boldsymbol{\eta}, \tilde{\mathbf{x}}) \simeq \ell_0(\boldsymbol{\theta}, \boldsymbol{\eta}, \tilde{\mathbf{x}}) - \delta \sum_{i=1}^{n}\sum_{j=1}^{m} i_{\eta_{ij}}^{-1/2}\Big\{ & I_{ij}B(y_{ij}, \boldsymbol{\theta}, \mathbf{x}_{ij})\frac{\partial}{\partial \eta_i}H_C(y_{ij}, \eta_{ij}) \\ & + (1 - I_{ij})(1 - Z_{ij})\frac{\partial}{\partial \eta_{ij}}H_C(y_{ij}, \eta_{ij})\mu(y_{ij}, \boldsymbol{\theta}, \mathbf{x}_{ij}) \\ & - Z_{ij}(1 - I_{ij})s_C(y_{ij}, \eta_{ij})\mu(y_{ij}, \boldsymbol{\theta}, \mathbf{x}_{ij})\Big\}, \end{aligned} \tag{4.18}$$

where

$$\mu(y_{ij}, \boldsymbol{\theta}, \mathbf{x}_{ij}) = \frac{\displaystyle\int_{y_{ij}}^{\infty} f_T(u, \boldsymbol{\theta}, \mathbf{x}_{ij})B(u, \boldsymbol{\theta}, \mathbf{x}_{ij})du}{S_T(y_{ij}, \boldsymbol{\theta}, \mathbf{x}_{ij})}.$$

The log-likelihood for the model where T and C are independent is given by

$$\begin{aligned}\ell_0(\boldsymbol{\theta}, \boldsymbol{\eta}, \tilde{\mathbf{x}}) = \sum_{i=1}^{n}\sum_{j=1}^{m}\Big\{ & I_{ij}\log h_T(y_{ij}, \boldsymbol{\theta}, \mathbf{x}_{ij}) + Z_{ij}(1 - I_{ij})\log h_C(y_{ij}, \eta_{ij}) \\ & - H_T(y_{ij}, \boldsymbol{\theta}, \mathbf{x}_{ij}) - H_C(y_{ij}, \eta_{ij})\Big\}. \end{aligned} \tag{4.19}$$

The derivation of a sensitivity analysis for $\boldsymbol{\theta}$ requires use of the Taylor expansions about $\boldsymbol{\theta}$ of the vector score functions

$$\boldsymbol{r}_0(\hat{\boldsymbol{\theta}}_0) = \frac{\partial}{\partial \boldsymbol{\theta}_0}\ell_0(\boldsymbol{\theta}, \boldsymbol{\eta}, \tilde{\mathbf{x}}) \quad \text{and} \quad \boldsymbol{r}_\delta(\hat{\boldsymbol{\theta}}_\delta) = \frac{\partial}{\partial \boldsymbol{\theta}_\delta}\ell_\delta(\boldsymbol{\theta}, \boldsymbol{\eta}, \tilde{\mathbf{x}})$$

to obtain an approximation to the change in the estimated parameters. The expansions ignoring any quadratic terms or higher are

$$\begin{aligned}\boldsymbol{r}_0(\hat{\boldsymbol{\theta}}_0) &\simeq \boldsymbol{r}_0(\boldsymbol{\theta}) - (\hat{\boldsymbol{\theta}}_0 - \boldsymbol{\theta})i(\boldsymbol{\theta}, \tilde{\mathbf{x}}) = 0, \text{and} \\ \boldsymbol{r}_\delta(\hat{\boldsymbol{\theta}}_\delta) &\simeq \boldsymbol{r}_\delta(\boldsymbol{\theta}) - (\hat{\boldsymbol{\theta}}_\delta - \boldsymbol{\theta})i(\boldsymbol{\theta}, \tilde{\mathbf{x}}) = 0. \end{aligned} \tag{4.20}$$

An expression for the difference in the vector of parameter estimates can be obtained by rearranging the linear expansions in (4.20). This is given by

$$\hat{\boldsymbol{\theta}}_\delta - \hat{\boldsymbol{\theta}}_0 \simeq i(\boldsymbol{\theta}, \tilde{\mathbf{x}})^{-1}(\boldsymbol{r}_\delta(\boldsymbol{\theta}) - \boldsymbol{r}_0(\boldsymbol{\theta})), \tag{4.21}$$

where the kth component of $\boldsymbol{r}_\delta(\boldsymbol{\theta}) - \boldsymbol{r}_0(\boldsymbol{\theta})$ is

$$\delta \sum_{i=1}^{n}\sum_{j=1}^{m} i_{\eta_{ij}}^{-1/2}\Big\{ Z_{ij}(1-I_{ij})s_C(y_{ij},\eta_{ij})\frac{\partial}{\partial\theta_k}\mu(y_{ij},\boldsymbol{\theta},\mathbf{x}_{ij}) \\ - (1-I_{ij})(1-Z_{ij})\frac{\partial}{\partial\eta_{ij}}H_C(y_{ij},\eta_{ij})\frac{\partial}{\partial\theta_k}\mu(y_{ij},\boldsymbol{\theta},\mathbf{x}_{ij}) \\ - I_{ij}\frac{\partial}{\partial\theta_k}B(y_{ij},\boldsymbol{\theta},\mathbf{x}_{ij})\frac{\partial}{\partial\eta_{ij}}H_C(y_{ij},\eta_{ij})\Big\}, \tag{4.22}$$

and the (k,l)th element of the information matrix $i(\boldsymbol{\theta},\tilde{\mathbf{x}})$ is

$$-\frac{\partial^2}{\partial\theta_k\partial\theta_l}\ell_0(\boldsymbol{\theta},\boldsymbol{\eta},\tilde{\mathbf{x}}).$$

Again (4.22) can be greatly simplified by assuming a proportional hazards structure. This is done using equations similar to those in Section 4.2.2, except that the derivative of the integrated hazard function for T is now

$$\frac{\partial}{\partial\theta_k}H_T(y_{ij},\boldsymbol{\theta},\mathbf{x}_{ij}) = x_{ijk}H_T(y_{ij},\boldsymbol{\theta},\mathbf{x}_{ij}).$$

So, the expression for the kth component of $\boldsymbol{r}_\delta(\boldsymbol{\theta}) - \boldsymbol{r}_0(\boldsymbol{\theta})$ becomes:

$$\delta \sum_{i=1}^{n}\sum_{j=1}^{m} x_{ijk}H_T(y_{ij},\boldsymbol{\theta},\mathbf{x}_{ij})\Big[H_C(y_{ij},\eta_{ij}) - (1-I_{ij})Z_{ij}\Big], \tag{4.23}$$

and the (k,l)th element of $i(\boldsymbol{\theta},\tilde{\mathbf{x}})$ is

$$\sum_{i=1}^{n}\sum_{j=1}^{m} x_{ijk}x_{ijl}H_T(y_{ij},\boldsymbol{\theta},\mathbf{x}_{ij}). \tag{4.24}$$

We can then use the parameter estimates found using the sensitivity analysis in (4.21) to obtain an approximation to the change in the linear predictors, $\hat{w}_\delta(\mathbf{x}_{ij})$ and $\hat{w}_0(\mathbf{x}_{ij})$ for the ith individual in the jth interval in the data set with covariate vector $\mathbf{x}_{ij}$. The equation used to do this is

$$\hat{w}_\delta(\mathbf{x}_{ij}) - \hat{w}_0(\mathbf{x}_{ij}) = \hat{\boldsymbol{\theta}}_\delta'\mathbf{x}_{ij} - \hat{\boldsymbol{\theta}}_0'\mathbf{x}_{ij}.$$

This approach that gives a sensitivity analysis for $\boldsymbol{\theta}$ is much more computationally time consuming carrying out the sensitivity analysis on $w(\mathbf{x})$. However, it is useful as it allows the effect of informative censoring on individual parameters to be estimated. Therefore, a sensitivity analysis for $\boldsymbol{\theta}$ will be applied to the Liver Registration data set using the expression in (4.21).

4.3 Application to the Liver Registration Data set

We will now apply the sensitivity analyses derived in Section 4.2 to the Liver Registration dataset. We assume that the lifetime and censoring variables each have piecewise exponential marginal distributions. Starting values for the cut points were chosen by splitting the time period into intervals with roughly equal numbers of observations in each interval. Then, the models with the intervals that give the largest value of the likelihood were found for 3 and 4 intervals. The log-cumulative hazard plots were then examined to check if the assumed model is appropriate. The 3-interval model was found to be appropriate and the 4-interval model did not seem to give any improvement. Therefore in the interest of parsimony, we used the 3-interval model with cut points at 40 and 165 days.

To determine whether this chosen model gives a significantly better fit than the corresponding standard Weibull models, the differences in $-2 \log \hat{L}$ for the models were found. If the true hazard is Weibull, then the difference in $-2 \log \hat{L}$ for the Weibull model and the piecewise exponential model should be approximately χ^2_{m-2}. The piecewise exponential model was significantly better than the Weibull for time to censoring ($p < 0.0001$) but not for time to death (p=0.85). As the same form of model must be used for both time to death and time to censoring when applying the sensitivity analysis, then the use of piecewise exponential models for the marginal distributions of the failure and censoring variables is justified for the Liver Registration data.

These models can be fitted using standard statistical software packages (such as PROC LIFEREG in SAS) as long as the data have been correctly formatted in a counting process format that gives both a start and stop time for the observation. There are multiple lines of data for an individual if they are at risk in multiple intervals. The exposure times that were defined earlier are used along with the indicator variables, I_{ij} and Z_{ij}.

4.3.1 Sensitivity analysis for scalar parameters

Firstly, the sensitivity analysis is applied to the Liver Registration data set assuming that T and C have piecewise exponential marginals distributions with scalar parameters in each interval. The derivation of this method was given in Section 4.2, with the simplified sensitivity analysis equation given in Section 4.2.2. The parameters of interest here are the scalar parameters for T in each interval. The scalar parameters for C in each interval are treated as nuisance parameters

When we assume that T and C have piecewise exponential marginals distributions

with scalar parameters in each interval, then the hazards and associated functions are

$$\begin{aligned} h_T(t_j,\theta_j) &= e^{\theta_j}, & h_C(c_j,\gamma_j) &= e^{\gamma_j}, \\ H_T(t_j,\theta_j) &= e^{\theta_j}t_j, & H_C(c_j,\gamma_j) &= e^{\gamma_j}c_j, \\ S_T(t_j,\theta_j) &= \exp(-e^{\theta_j}t_j), \quad \text{and} & S_C(c_j,\gamma_j) &= \exp(-e^{\gamma_j}c_j). \end{aligned} \tag{4.25}$$

The form of the functions in (4.25) can be substituted into (4.14) to give the sensitivity analysis equation

$$\theta_{\delta j} - \theta_{0j} \simeq \delta \frac{\sum_{i\in R_j}\{e^{\hat{\gamma}_j}y_{ij}^2 - y_{ij}(1-I_{ij})Z_{ij}\}}{\sum_{i\in R_j} y_{ij}}, \tag{4.26}$$

for the jth interval. We can see that (4.26) has no dependence on the parameter of interest in the jth interval, θ_j, but does require a value of the nuisance parameter γ_j to be used. The value of γ_j that will be used is the estimate found using the likelihood in (4.5), which assumes non-informative censoring. The maximum likelihood estimate of γ_j is

$$\hat{\gamma}_j = \frac{\sum_{i\in R_j} Z_{ij}(1-I_{ij})}{\sum_{i\in R_j} y_{ij}}.$$

The values of $\hat{\gamma}_j$ found for the Liver Registration data set are -4.9189, -5.2320 and -5.5695 for intervals 1, 2 and 3 respectively.

Table 4.1 shows the approximate values of $\hat{\theta}_{\delta j} - \hat{\theta}_{0j}$ found by applying the sensitivity analysis equation in (4.26). We can see that the sensitivity analysis suggests that the largest changes in parameter estimates occurs in the final interval for the Liver Registration data set.

Interval (j)	$\hat{\theta}_{0.2j} - \hat{\theta}_{0j}$	$\hat{\theta}_{0.3j} - \hat{\theta}_{0j}$
1	0.0274	0.0411
2	0.0607	0.0911
3	0.1343	0.2015

Table 4.1: Table showing the estimated change in the parameter estimates for each interval from the sensitivity analysis using $\delta = 0.2$ and $\delta = 0.3$.

Table 4.2 shows the parameter estimates for time to death assuming non-informative censoring, along with the approximate parameter estimates for $\delta = 0.2$ and $\delta = 0.3$. The

values $\hat{\theta}_{0j}$ are the maximum likelihood estimates given by

$$\frac{\sum_{i \in R_j} I_{ij}}{\sum_{i \in R_j} y_{ij}}.$$

The values $\hat{\theta}_{0.2j}$ and $\hat{\theta}_{0.3j}$ are the approximate parameter estimates found by adding the values from the sensitivity analysis in Table 4.1 to $\hat{\theta}_{0j}$.

k	$\hat{\theta}_{0k}$	p-value	$\hat{\theta}_{0.2k}$	p-value	$\hat{\theta}_{0.3k}$	p-value
1	-6.7799	< 0.0001	-6.7525	< 0.0001	-6.7387	< 0.0001
2	-6.9056	< 0.0001	-6.8448	< 0.0001	-6.8145	< 0.0001
3	-7.6375	< 0.0001	-7.5031	< 0.0001	-7.4360	< 0.0001

Table 4.2: Table showing the parameter estimates for time to death assuming non-informative censoring, along with the approximate parameter estimates for $\delta = 0.2$ and $\delta = 0.3$ found using the results in Table 4.1.

4.3.2 Sensitivity analysis including covariates

We will now apply sensitivity analyses that include covariates to the Liver Registration data set. There are two methods of doing this, either a sensitivity analysis for $w(\mathbf{x})$ or a sensitivity analysis for $\boldsymbol{\theta}$, both of which are detailed in Section 4.2.3. We shall apply both methods to the data set so that the results from each can be compared. The sensitivity analysis for $w(\mathbf{x})$ is applied first, followed by the sensitivity analysis for $\boldsymbol{\theta}$.

In the initial data analysis for the Liver Registration data set, it was found that primary liver disease category, recipient ethnicity, age and UKELD score are significant for time to death and primary liver disease category, UKELD score, recipient height and recipient blood group are significant for time to censoring. Therefore, these covariates should be included in the models used in the sensitivity analysis.

If we let $w_j(\mathbf{x}_i) = \boldsymbol{\theta}_j' \mathbf{x}_i$ and $z_j(\mathbf{x}_i) = \boldsymbol{\gamma}_j' \mathbf{x}_i$, then the hazards and associated functions for T and C with piecewise exponential marginal distributions can be expressed as:

$$\begin{aligned}
h_T(t_j, \boldsymbol{\theta}_j, \mathbf{x}_i) &= e^{w_j(\mathbf{x}_i)} & h_C(c_j, \boldsymbol{\gamma}_j, \mathbf{x}_i) &= e^{z_j(\mathbf{x}_i)} \\
H_T(t_j, \boldsymbol{\theta}_j, \mathbf{x}_i) &= e^{w_j(\mathbf{x}_i)} t_j & H_C(c_j, \boldsymbol{\gamma}_j, \mathbf{x}_i) &= e^{z_j(\mathbf{x}_i)} c_j \\
S_T(t_j, \boldsymbol{\theta}_j, \mathbf{x}_i) &= \exp(-e^{w_j(\mathbf{x}_i)} t_j) & S_C(c_j, \boldsymbol{\gamma}_j, \mathbf{x}_i) &= \exp(-e^{z_j(\mathbf{x}_i)} c_j)
\end{aligned} \tag{4.27}$$

These can now be substituted into (4.17) to approximate $\hat{w}_{\delta j}(\mathbf{x}) - \hat{w}_{0j}(\mathbf{x})$. To calculate this we need $\hat{z}_{0j}(\mathbf{x})$, which is the estimated linear predictor for time to censoring assuming

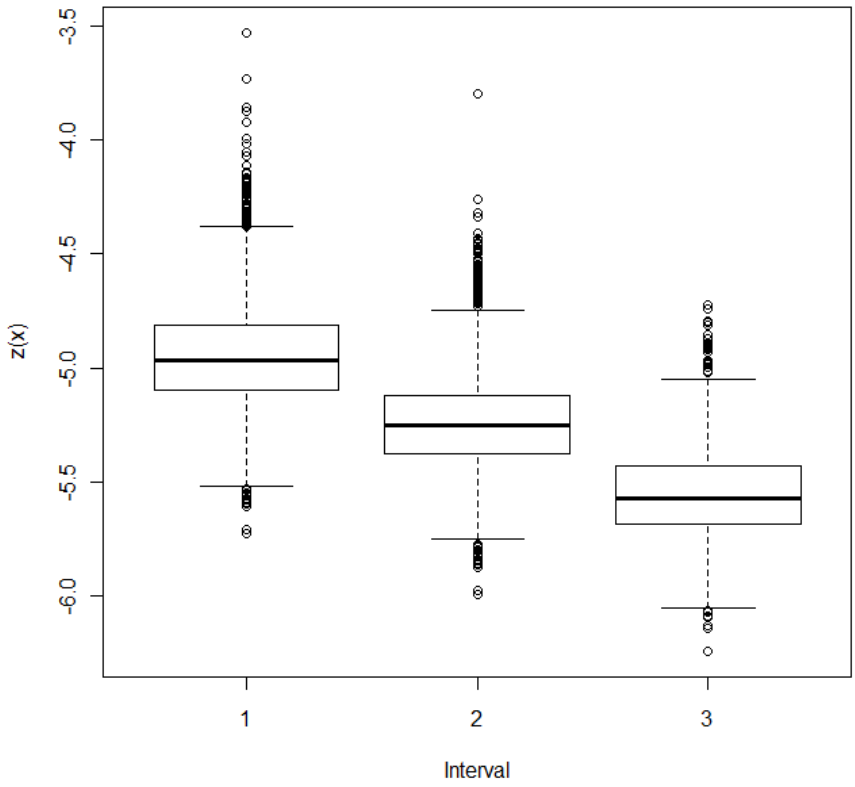

Figure 4.1: Boxplots showing the distribution of $\hat{z}_{0j}(\mathbf{x})$ in each of the three intervals for the Liver Registration data.

non-informative censoring. The distributions of $\hat{z}_{0j}(\mathbf{x})$ for the Liver Registration data are shown by the boxplots in Figure 4.1. We see that the median value of $\hat{z}_{0j}(\mathbf{x})$ decreases across the intervals, which shows that the hazard of censoring is generally smaller in the later intervals. We also see that the majority of patients have values of $\hat{z}_{0j}(\mathbf{x})$ that fall in the middle of the observed range for each interval, with only a small number at either of the extremes.

The approximation for $\hat{w}_{\delta j}(\mathbf{x}) - \hat{w}_{0j}(\mathbf{x})$ when conducting a sensitivity analysis on $w(\mathbf{x})$ is obtained by substituting the functions from (4.27) into (4.17). This then gives:

$$\hat{w}_{\delta j}(\mathbf{x}) - \hat{w}_{0j}(\mathbf{x}) \simeq \delta \frac{\sum_{i \in R_j} \{e^{\hat{z}_0(\mathbf{x})} y_{ij}^2 - y_{ij}(1 - I_{ij}) Z_{ij}\}}{\sum_{i \in R_j} y_{ij}}. \tag{4.28}$$

The equation in (4.28) requires the same covariate vector to be used in both the models for time to death and time to censoring. Therefore, we include age, recipient ethnicity,

primary liver disease category, UKELD score, recipient height and recipient blood group as covariates in the models for time to death and time to censoring.

The sensitivity analysis in (4.28) only considers an arbitrary vector of covariates $\mathbf{x}$, when we want to assess the change in parameter estimates for all individuals in the dataset. Therefore we need to plot the estimated value of $\hat{w}_\delta(\mathbf{x}) - \hat{w}_0(\mathbf{x})$ against the entire range of values that $\hat{z}_{0j}(\mathbf{x})$ takes across all the individuals in jth interval, which is shown for all 3 intervals in Figure 4.2. This figure shows the plot for $\delta = 0.2$ and 0.3. It can be seen from Figure 4.2 that the second and third intervals have larger estimated values of $\hat{w}_{\delta j}(\mathbf{x}) - \hat{w}_{0j}(\mathbf{x})$ than the first interval. The largest values of $\hat{w}_{\delta j}(\mathbf{x}) - \hat{w}_{0j}(\mathbf{x})$ are observed for the patients with the largest values of $\hat{z}_{0j}(\mathbf{x})$ or the highest hazards of censoring. However if we consider the distributions of $\hat{z}_{0j}(\mathbf{x})$ shown in Figure 4.1, then we can see that only a small number of individuals will have these large changes in $\hat{w}_{\delta j}(\mathbf{x}) - \hat{w}_{0j}(\mathbf{x})$. This means the effect of informative censoring is small for the majority of patients in the Liver Registration data. However, as some individuals have a large estimated change in the linear predictors, then any inferences may be misleading if non-informative censoring was assumed, and there is even a moderate amount of dependence between the time to death and time to censoring variables.

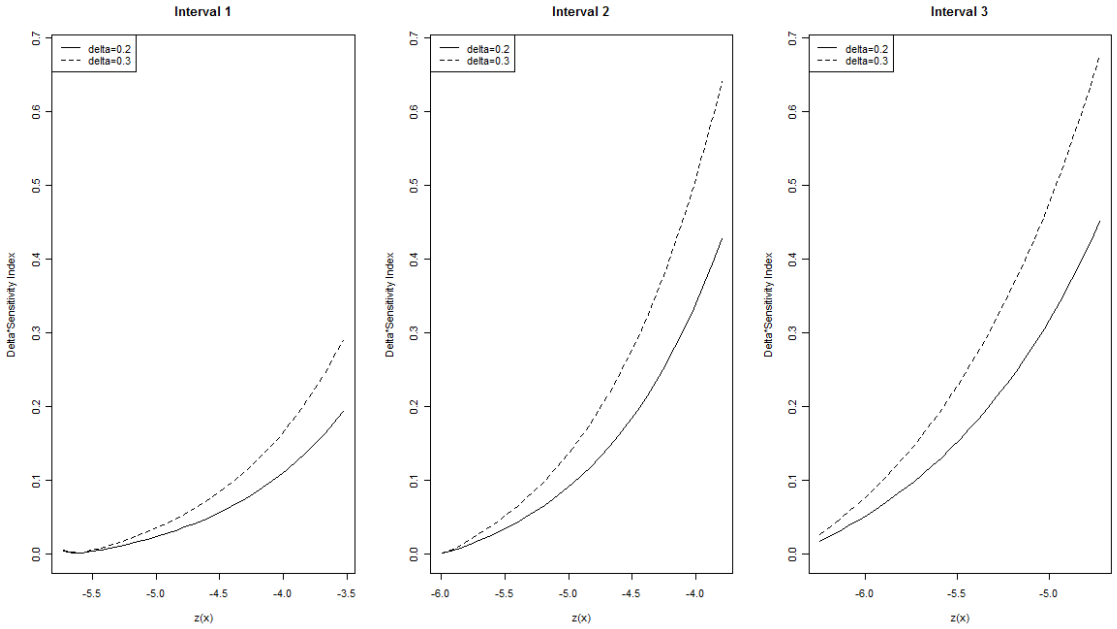

Figure 4.2: Plot of sensitivity analysis expression in (4.28) for observed values of $\hat{z}_{0j}(\mathbf{x})$ for the Liver Registration data in each of the three intervals with $\delta = 0.2$, 0.3, when applying the sensitivity analysis for $w(\mathbf{x})$

Now the sensitivity analysis for $\boldsymbol{\theta}$ is applied to the Liver Registration data set. This

allows the change in individual components of the vector of parameter estimates to be estimated, instead of just looking at the change in the linear predictor $w(\mathbf{x})$. The equations for this sensitivity analysis were also derived in Section 4.2.3.

The hazards and associated functions for T and C are now given by

$$\begin{aligned} h_T(t_j, \boldsymbol{\theta}, \mathbf{x}_{ij}) &= e^{\boldsymbol{\theta}'\mathbf{x}_{ij}}, & h_C(c_j, \boldsymbol{\gamma}, \mathbf{x}_{ij}) &= e^{z(\mathbf{x}_{ij})}, \\ H_T(t_j, \boldsymbol{\theta}, \mathbf{x}_{ij}) &= e^{\boldsymbol{\theta}'\mathbf{x}_{ij}} t_j & H_C(c_j, \boldsymbol{\gamma}, \mathbf{x}_{ij}) &= e^{z(\mathbf{x}_{ij})} c_j, \\ S_T(t_j, \boldsymbol{\theta}, \mathbf{x}_{ij}) &= \exp(-e^{\boldsymbol{\theta}'\mathbf{x}_{ij}} t_j), \quad \text{and} & S_C(c_j, \boldsymbol{\gamma}, \mathbf{x}_{ij}) &= \exp(-e^{z(\mathbf{x}_{ij})} c_j), \end{aligned} \tag{4.29}$$

where the value of interest is the parameter vector $\boldsymbol{\theta}$, with $z(\mathbf{x})$ again being treated as a nuisance parameter. The expressions in (4.29) can be substituted in (4.23) from Section 4.2.3 to give

$$\delta \sum_{i=1}^{n} \sum_{j=1}^{m} \{x_{ijk} e^{\boldsymbol{\theta}'\mathbf{x}_{ij}} y_{ij} [e^{z(\mathbf{x}_{ij})} y_{ij} - (1 - I_{ij}) Z_{ij}]\}, \tag{4.30}$$

for the kth component of $\boldsymbol{r}_\delta(\boldsymbol{\theta}) - \boldsymbol{r}_0(\boldsymbol{\theta})$. The (k, l)th element of the information matrix $i(\boldsymbol{\theta}, \tilde{\mathbf{x}})$ also becomes

$$\sum_{i=1}^{n} \sum_{j=1}^{m} x_{ijk} x_{ijl} e^{\boldsymbol{\theta}'\mathbf{x}_{ij}} y_{ij} \tag{4.31}$$

when the form of the integrated hazard function for T in (4.29) is substituted into (4.24) in Section 4.2.3. The expressions in (4.30) and (4.31) can then be used in

$$\hat{\boldsymbol{\theta}}_\delta - \hat{\boldsymbol{\theta}}_0 \simeq i(\boldsymbol{\theta}, \tilde{\mathbf{x}})^{-1} (\boldsymbol{r}_\delta(\boldsymbol{\theta}) - \boldsymbol{r}_0(\boldsymbol{\theta})),$$

to conduct a sensitivity analysis for $\boldsymbol{\theta}$.

We can see that, unlike the previous expressions for the sensitivity analyses for scalar parameters and $w(\mathbf{x})$, (4.30) and (4.31) contain the parameter vector of interest, $\boldsymbol{\theta}$, as well as the nuisance parameter, $z(\mathbf{x})$. This means that values of $\boldsymbol{\theta}$ will need to substituted into (4.30) and (4.31) along with values of $z(\mathbf{x})$ to carry out the sensitivity analysis. The values used are the MLEs from the piecewise exponential model assuming non-informative censoring.

Also, although it is assumed, for notational simplicity, that the models for T and C use the same covariate vector $\mathbf{x_{ij}}$, it is possible to use separate models for the two variables. Therefore, primary liver disease category, recipient ethnicity, recipient age and UKELD score will be included in the model for time to death and primary liver disease category, UKELD score, recipient height and recipient blood group in the model for time to censoring.

Table 4.3 shows the estimated values of the components of $\hat{\boldsymbol{\theta}}_\delta - \hat{\boldsymbol{\theta}}_0$ for $\delta = 0.2$ and $\delta = 0.3$. We see that for some covariates there are positive changes in the parameter estimates, while others have negative changes in the parameter estimates. Positive values in Table

4.3 mean that the element of $\hat{\boldsymbol{\theta}}_\delta$ for that covariate is larger than the corresponding element of $\hat{\boldsymbol{\theta}}_0$. So, this suggests that the hazard ratio of the covariate is being underestimated by the model assuming non-informative censoring. Conversely, negative values in Table 4.3 mean that the parameter estimate for the covariate from the model assuming informative censoring is smaller than the corresponding parameter estimate from the model assuming non-informative censoring. Therefore, the sensitivity analysis is suggesting that the hazard ratio for these covariates are overestimated by the model that assumes $\delta = 0$.

The sensitivity analysis for $\boldsymbol{\theta}$ suggests that the majority of the hazard ratios for the levels of primary liver disease category are overestimated by the model that assumes non-informative censoring, apart from patients with hepatitis B infection, cancer or metabolic liver disease, whose hazard ratios are being underestimated. Similarly, most of the hazard ratios for the levels of recipient ethnicity are being overestimated with the exception of patients of oriental ethnic origin, whose hazard ratio is being underestimated. The sensitivity analysis also suggests that there should be small alterations made to the parameter estimates for the UKELD score and recipient age from the model assuming non-informative censoring. However, the parameter estimate for recipient age should be reduced while the parameter estimate for the UKELD score needs to be increased.

Table 4.4 shows the approximate parameter estimates for piecewise exponential models assuming $\delta = 0.2$ and $\delta = 0.3$ respectively. The parameter estimates for the model assuming non-informative censoring are also shown. These parameter estimates assuming $\delta = 0.2$ and $\delta = 0.3$ are obtained by adding the values of $\hat{\boldsymbol{\theta}}$ in Table 4.4 to the values in Table 4.3. The p-values of all the estimates are also shown. These are calculated using the standard errors of the estimates from the model assuming non-informative censoring. This can be done as Siannis et al. (2005) show that

$$\{\mathrm{Var}(\hat{\theta}_\delta)\}^{1/2} \simeq \{\mathrm{Var}(\hat{\theta}_0)\}^{1/2} + O(\delta^2).$$

Only linear values of δ are considered in the sensitivity analysis so the standard error of the parameter estimate from the model assuming informative censoring can be approximated by the standard error of the parameter estimate from the model assuming non-informative censoring. This approximation should only be used if the value of δ is fairly small.

The approximate values of $\hat{\boldsymbol{\theta}}_{0.2}$ and $\hat{\boldsymbol{\theta}}_{0.3}$ given in Table 4.4 can be used to find the change in the estimated linear predictor for T under this sensitivity analysis. This is done for each individual in the data set using the expression

$$\hat{w}_\delta(\mathbf{x}_{ij}) - \hat{w}_0(\mathbf{x}_{ij}) = \hat{\boldsymbol{\theta}}'_\delta \mathbf{x}_{ij} - \hat{\boldsymbol{\theta}}'_0 \mathbf{x}_{ij}.$$

The largest value of this change that is estimated by the sensitivity analysis is 0.2289 for $\delta = 0.2$ and 0.3434 for $\delta = 0.3$. These values are much smaller than the corresponding

Parameter	$\hat{\boldsymbol{\theta}}_{0.2} - \hat{\boldsymbol{\theta}}_0$	$\hat{\boldsymbol{\theta}}_{0.3} - \hat{\boldsymbol{\theta}}_0$
Intercept	0.18243	0.27364
PLD - PBC	-0.03830	-0.05746
PLD - PSC	-0.01971	-0.02956
PLD - ALD	-0.02469	-0.03703
PLD - AID	-0.02944	-0.04416
PLD - HCV	-0.01639	-0.02458
PLD - HBV	0.02421	0.03631
PLD - Cancer	0.04255	0.06383
PLD - Metabolic	0.01421	0.02132
PLD - Other	-0.04777	-0.07165
Ethnicity - White	-0.02683	-0.04025
Ethnicity - Asian	-0.00012	-0.00018
Ethnicity - Black	-0.03322	-0.04983
Ethnicity - Chinese	0.00443	0.00665
UKELD	0.00020	0.00029
Age	-0.00013	-0.00020
j - Interval 1	-0.10709	-0.16063
j - Interval 2	-0.08191	-0.12286

Table 4.3: Table showing the components of $\hat{\boldsymbol{\theta}}_\delta - \hat{\boldsymbol{\theta}}_0$ approximated by the sensitivity analysis for $\delta = 0.2$ and $\delta = 0.3$.

Parameter	$\hat{\boldsymbol{\theta}}_0$	p-value	$\hat{\boldsymbol{\theta}}_{0.2}$	p-value	$\hat{\boldsymbol{\theta}}_{0.3}$	p-value
Intercept	-20.54993	< 0.0001	-20.36750	< 0.0001	-20.27629	< 0.0001
PLD - PBC	-0.23181	0.49682	-0.27011	0.42849	-0.28926	0.39649
PLD - PSC	-0.93303	0.01862	-0.95274	0.01627	-0.96259	0.01520
PLD - ALD	-0.46799	0.13534	-0.49268	0.11592	-0.50502	0.10707
PLD - AID	-0.02429	0.94191	-0.05373	0.87194	-0.06845	0.83729
PLD - HCV	0.23191	0.48051	0.21553	0.51206	0.20733	0.52823
PLD - HBV	-0.44046	0.44924	-0.41626	0.47455	-0.40415	0.48749
PLD - Cancer	-1.46458	0.05627	-1.42203	0.06381	-1.40075	0.06789
PLD - Metabolic	0.64451	0.07151	0.65872	0.06548	0.66583	0.06262
PLD - Other	0.36075	0.28898	0.31298	0.35759	0.28910	0.39545
Ethnicity - White	0.97872	0.33155	0.95189	0.34498	0.93848	0.35182
Ethnicity - Asian	-0.02734	0.97917	-0.02746	0.97908	-0.02752	0.97904
Ethnicity - Black	0.92243	0.41008	0.88921	0.42715	0.87260	0.43583
Ethnicity - Chinese	-0.72652	0.61251	-0.72209	0.61468	-0.71987	0.61577
UKELD	0.19145	< 0.0001	0.19164	< 0.0001	0.19174	< 0.0001
Age	0.03019	< 0.0001	0.03005	< 0.0001	0.02999	< 0.0001
j - Interval 1	0.21283	0.22803	0.10574	0.54923	0.05220	0.76750
j - Interval 2	0.47753	0.00265	0.39562	0.01279	0.35467	0.02562

Table 4.4: Table showing the parameter estimates for the model assuming non-informative censoring, along with the parameter estimates approximated by the sensitivity analysis for $\delta = 0.2$ and $\delta = 0.3$. The p-values are also shown, these are all found using the standard errors from the model assuming non-informative censoring.

values estimated by the sensitivity analysis for $w(\mathbf{x})$. Therefore an investigation into which of these two sensitivity methods is more accurate would be useful.

4.4 Summary

In this chapter, we present a general method that allows us to estimate the change in parameter estimates for piecewise parametric models if we assume a small amount of informative censoring instead of non-informative censoring. The method is first derived assuming only scalar parameters in each interval in the models for time to death and time to censoring. It is then extended to include a vector of covariates. To include covariates we need to use piecewise parametric models that can be expressed in terms of parameter vectors that remain constant over all intervals, as the parameter estimates in all the intervals need to be estimated at the same time.

The method presented in this chapter is a compromise between the sensitivity analysis given in Siannis et al. (2005) and Siannis (2004) and the sensitivity analysis in Siannis (2011). Our method has the flexibility of the Cox model that is used in Siannis (2011), but it computationally simpler like the methods in Siannis et al. (2005) and Siannis (2004).

When including covariates in the method, it is possible to apply a sensitivity analysis for either a linear predictor or a vector of parameters. The sensitivity analysis for a linear predictor is computationally simpler but the sensitivity analysis for a parameter vector allows us to examine the effect on individual parameter estimates not just the overall effect. These two methods give very different values of the estimated changes in the parameter estimates, therefore an investigation into which method is more accurate would be useful. This is why the model that assumes informative censoring that can be approximated using the sensitivity analysis will be fitted to the Liver Registration data set in Chapter 5.

Chapter 5

Evaluating and Extending Sensitivity Analysis

Having developed and applied the sensitivity analysis, we now investigate its properties. Firstly, we will assess how close the approximation is for the dataset of interest. However, we also want to be able draw some more general conclusions about the behaviour of the sensitivity analysis. Therefore, simulations will be used to test the accuracy of the sensitivity analysis for many different combinations of the parameters. Based on the results of these simulations, we will make recommendations on possible ways to improve the sensitivity analysis.

To find the true difference between the parameter estimates, it will be necessary to fit the dependence model that does not approximate the form of the joint distribution. Thus we consider a slightly different change in the parameters than is approximated by our sensitivity analysis equation. We do this because it is more useful to know how accurate the sensitivity analysis is at estimating the change in parameter estimates from the independence model to the dependence model that does not make any simplifying assumptions. These assumptions were necessary in the last chapter to get a closed form of the likelihood to work with. The fitting of the dependence model will be detailed below, including a brief description of some of the numerical methods that need to be used to fit the model.

The dependence model is fitted to our example dataset and compared to the results obtained for the sensitivity analysis in the previous chapter. Then a simulation study will be used to assess the accuracy of the sensitivity analysis in a variety of situations. All these investigations will consider the alternative sensitivity analysis as it is easier to apply when using piecewise exponential models. Finally, a possible way of improving the sensitivity analysis to overcome some of the issues that are raised by the simulation study will be outlined.

5.1 Fitting dependence model

We write the joint density of T and C_I, as

$$f_{T,C}(t_j, c_j) = f_T(t_j, \theta_j) f_{C|T}(c_j | t_j, \gamma_j, \delta, \theta_j) \tag{5.1}$$

As in Section 4.1 we assume that

$$f_{C|T}(c_j | t_j, \gamma_j, \delta, \theta_j) = f_C(c_j, \gamma_j + \delta i_{\gamma_j}^{-1/2} B(t_j, \theta_j)), \tag{5.2}$$

with $i_{\gamma_j} = 1$ and $B(t_j, \theta_j) = i_{\theta_j} s_T(t_j, \theta_j) = 1 - e^{\theta_j} t$ under our proportional hazards structure, given in Section 4.2.2.

In addition, we assume piecewise exponential marginal models for both T and C_I, so

$$f_T(t_j, \boldsymbol{\theta}, \mathbf{x}_j) = e^{\boldsymbol{\theta}' \mathbf{x}_j} e^{-\exp(\boldsymbol{\theta}' \mathbf{x}_j) t} \quad \text{and} \quad f_C(c, \eta) = e^{\eta_j} e^{-\exp(\eta_j) c},$$

where the linear combination $\boldsymbol{\gamma}' \mathbf{x}_j$ is replaced by a scalar parameter η_j, as it is just a nuisance parameter. The vector $\boldsymbol{\theta}$ here is set up in the same way as the vector of parameters in Section 4.2.3. It has a common intercept for all the intervals which is adjusted for each interval by a factor giving the contrast between the baseline intercept and the interval under consideration.

If we combine (5.1) and (5.2), and then substitute the exponential forms into the resulting equation, then we obtain:

$$f_{T,C}(t_j, c_j) = e^{\boldsymbol{\theta}' \mathbf{x}_j} e^{-\exp\{\boldsymbol{\theta}' \mathbf{x}_j\} t_j} e^{\eta_j + \delta(1 - \exp\{\boldsymbol{\theta}' \mathbf{x}_j\} t_j)} e^{-\exp(\eta_j + \delta(1 - \exp\{\boldsymbol{\theta}' \mathbf{x}\} t_j)) c_j}. \tag{5.3}$$

The parameter estimates for the full model will be obtained by finding the maximum likelihood estimates of the likelihood detailed below

$$\begin{aligned} \ell_\delta(\boldsymbol{\theta}, \boldsymbol{\gamma}, \tilde{\mathbf{x}}) = \sum_{i=1}^{n} \sum_{j=1}^{m} \Big\{ & I_{ij} \log K_1(y_{ij}) + Z_{ij}(1 - I_{ij}) \log K_2(y_{ij}) \\ & + (1 - I_{ij})(1 - Z_{ij}) \log K_3(y_{ij}) \Big\}, \end{aligned} \tag{5.4}$$

where

$$K_1(y_{ij}) = e^{\boldsymbol{\theta}' \mathbf{x}_{ij}} e^{-\exp\{\boldsymbol{\theta}' \mathbf{x}_{ij}\} y_{ij}} e^{-\exp(\eta_j + \delta(1 - \exp\{\boldsymbol{\theta}' \mathbf{x}_{ij}\} y_{ij})) y_{ij}}$$

$$K_2(y_{ij}) = \int_{y_{ij}}^{\infty} e^{\boldsymbol{\theta}' \mathbf{x}_{ij}} e^{-\exp\{\boldsymbol{\theta}' \mathbf{x}_{ij}\} u} e^{\eta_j + \delta(1 - \exp\{\boldsymbol{\theta}' \mathbf{x}_{ij}\} u)} e^{-\exp(\eta_j + \delta(1 - \exp\{\boldsymbol{\theta}' \mathbf{x}_{ij}\} u)) y_{ij}} du$$

and

$$K_3(y_{ij}) = \int_{y_{ij}}^{\infty} e^{\boldsymbol{\theta}' \mathbf{x}_{ij}} e^{-\exp\{\boldsymbol{\theta}' \mathbf{x}\} u} e^{-\exp(\eta_j + \delta(1 - \exp\{\boldsymbol{\theta}' \mathbf{x}_{ij}\} u)) y_{ij}} du. \tag{5.5}$$

These were obtained by substituting the form of the joint distribution given in (5.3) into (4.2). From now on we will define $\hat{\boldsymbol{\theta}}_\delta$ as the vector of values that maximises the likelihood

in (5.4). This should be close to the value of $\hat{\boldsymbol{\theta}}_\delta$ defined in the previous chapter if the approximation used in (4.3) is a good approximation of the joint distribution in (5.3).

The two integrals in (5.5) cannot be evaluated analytically. We use Gauss-Laguerre quadrature as it will be easy to transform the integrals in (5.5) into the form

$$\int_0^\infty e^{-y} g(y) dy$$

so we can approximate the integral by

$$\sum_{j=1}^{N} w_j g(v_j),$$

where w_j and v_j are respectively the set of weights and abscissas for the integer N. Here $N = 32$ is used. We can then find the maximum likelihood estimates using the downhill simplex method of Nelder and Mead. This is inefficient but robust for functions where we can compute function evaluations but not derivatives. Both of these methods are outlined in greater detail in Press et al.(1992).

5.1.1 Fitting the dependence model to the Liver Registration data set

To be able to assess the accuracy of the sensitivity analysis developed in Chapter 4, it is necessary to fit the dependence model described in Section 5.1 to the Liver Registration data set. Firstly, this will be done assuming only scalar parameters in each interval for the model for time to death and time to censoring. This will allow the accuracy of the sensitivity analysis applied in Section 4.3.1 to be assessed. Covariates will then be included to allow the accuracy of the sensitivity analyses applied in Section 4.3.2 to be assessed. This should indicate whether a sensitivity analysis for the linear predictor $w(\mathbf{x})$ or a sensitivity analysis for the parameter vector $\boldsymbol{\theta}$ is more accurate.

The parameter estimates obtained by fitting the dependence model to the Liver Registration data set with scalar parameters in each interval for the model for time to death and time to censoring are given in Table 5.1. The parameter estimates obtained by fitting the corresponding independence model are included for comparison.

The sensitivity analysis from Section 4.3.1 which uses scalar parameters will now be reapplied using $\delta = 0.2698$, which is the fitted value from the dependence model. The results of this sensitivity analysis are given in Table 5.2. The estimated values of $\hat{\theta}_{0.2698k} - \hat{\theta}_{0k}$ found using the sensitivity analysis are compared to the observed values of $\hat{\theta}_{0.2698k} - \hat{\theta}_{0k}$ found by taking the difference of the values in Table 5.1. We can see that the sensitivity analysis overestimates the change in the parameter estimates for the first interval, but underestimates the change in the parameter estimates in the second and third intervals.

The dependence model including covariates will now be fitted to the Liver Registration data set. The explanatory variables for time to death used are age at registration, recipient

Parameter	Estimate from the independence model	Estimate from the dependence model
δ	-	0.2698
θ_1	-6.7799	-6.7571
θ_2	-6.9056	-6.8206
θ_3	-7.6375	-7.4458
γ_1	-4.9189	-4.9137
γ_2	-5.2320	-5.1008
γ_3	-5.5695	-5.2806

Table 5.1: The parameter estimates obtained by fitting the dependence model to the Liver registration data set assuming scalar parameters in each interval for the models for time to death and time to censoring. The parameter estimates from the independence model are also given for comparison.

k	Estimated value of $\hat{\theta}_{0.2698k} - \hat{\theta}_{0k}$	Observed value of $\hat{\theta}_{0.2698k} - \hat{\theta}_{0k}$
1	0.0370	0.0228
2	0.0819	0.0850
3	0.1812	0.1917

Table 5.2: The estimated values of $\hat{\theta}_{0.2698k} - \hat{\theta}_{0k}$ found using the sensitivity analysis from 4.3.1 and the observed values of $\hat{\theta}_{0.2698k} - \hat{\theta}_{0k}$ found using the values in Table 5.1.

ethnicity, primary liver disease category and UKELD score at registration. These are the same covariates included in the model for time to death in Section 4.3.2 when applying the sensitivity analysis for $\boldsymbol{\theta}$. However, only a scalar parameter is used for the model for time to censoring. This is because the model is already has a fairly large number of dimensions and including covariates for time to censoring would add enough extra dimensions to make the convergence of the algorithms in Section 5.1 too slow. This means that the results from this dependence model cannot be compared directly to the sensitivity analyses in Section 4.3.2, as they use covariates in their models for time to censoring.

The estimates for the dependence model obtained are given in Table 5.3. To see how much these vary from the estimates given by fitting the independence model, $\hat{\boldsymbol{\theta}}_0$ is also included in Table 5.3. The sensitivity analyses for $w(\mathbf{x})$ and $\boldsymbol{\theta}$ are carried out using $\delta = 0.2769$, which is the fitted value from the dependence model. This allows the direct comparison of the results of the sensitivity analyses with the results of the dependence model in Table 5.3.

The sensitivity analysis for $w(\mathbf{x})$ requires the same vector of covariates to be used in the model for time to death and the model for time to censoring. This means that the sensitivity analysis that is used for comparison with the results of the fitted dependence model includes primary liver disease category, recipient age, recipient ethnicity and UKELD score as covariates in the models for time to death and time to censoring. The plot in Figure 5.1 shows the estimated change in the linear predictor over the range of $\hat{z}_{0j}(\mathbf{x})$ observed in each of the intervals for the data for several values of δ. The solid line is the sensitivity analysis for $\delta = 0.2769$, which is the fitted value from the dependence model. The dashed lines are the sensitivity analyses for $\delta = 0.1377$ and $\delta = 0.4162$ which are the limits of 95% confidence interval for δ given in Table 5.3. These are included to show how the change in estimated linear predictor is greatly affected by the value of δ used.

The maximum change in the linear predictor estimated by the sensitivity analysis using $\delta = 0.2769$ is 0.6248, but the dashed lines suggest that this change could be anywhere between 0.3107 and 0.9391. However, when calculating the difference between $\hat{w}_\delta(\mathbf{x})$ and $\hat{w}_0(\mathbf{x})$ using the parameter estimates in Table 5.3, the largest difference observed was 0.3868.

This result shows that for the Liver Registration data, the sensitivity analysis tends to overestimate the change in the estimated linear predictors. However, only a small number of the patients in the data will have a discrepancy that is large. We already know that the sensitivity analysis gives the largest changes in $\hat{w}_\delta(\mathbf{x})$ and $\hat{w}_0(\mathbf{x})$ for the patients with the largest values of $\hat{z}_{0j}(\mathbf{x})$. From Figure 4.1, we know that only a small number of patients have values of $\hat{z}_{0j}(\mathbf{x})$ that are that large. So, for the majority of individuals in the Liver

Parameter	Estimate from independence model	Estimate from dependence model	Standard Error	95% Confidence Interval
δ	-	0.2769	0.0711	(0.1377,0.4162)
η	-	-5.0981	0.0221	(-5.1415,-5.0548)
θ Intercept	-20.5499	-20.0591	1.2500	(-22.5092,-17.6090)
Age	0.0302	0.0298	0.0061	(0.0180,0.0417)
Ethnicity - White	0.9787	1.0762	1.0105	(-0.9044,3.0568)
Ethnicity - Asian	-0.0273	0.0598	1.0495	(-1.9972,2.1169)
Ethnicity - Black	0.9224	0.9775	1.1223	(-1.2221,3.1771)
Ethnicity - Chinese	-0.7265	-0.5046	1.4260	(-3.2994,2.2903)
Ethnicity - Other	0	0		
PLD - PBC	-0.2318	-0.2548	0.3363	(-0.9140,0.4045)
PLD - PSC	-0.9330	-0.9391	0.3927	(-1.7089,-0.1694)
PLD - ALD	-0.4680	-0.4741	0.3083	(-1.0785,0.1302)
PLD - AID	-0.0243	-0.0693	0.3288	(-0.7137,0.5751)
PLD - HCV	0.2319	0.1965	0.3248	(-0.4401,0.8331)
PLD - HBV	-0.4405	-0.4647	0.5793	(-1.6001,0.6707)
PLD - Cancer	-1.4646	-1.5034	0.7639	(-3.0007,-0.0062)
PLD - Metabolic	0.6445	0.6279	0.3518	(-0.0616,1.3174)
PLD - Other	0.3608	0.2560	0.3357	(-0.4019,0.9139)
PLD - Acute	0	0		
UKELD	0.1914	0.1858	0.0099	(0.1664,0.2053)
j - Interval 1	0.2128	0.0241	0.1755	(-0.3198,0.3679)
j - Interval 2	0.4775	0.3317	0.1575	(0.0231,0.6403)
j - Interval 3	0	0		

Table 5.3: Parameter estimates, standard errors and 95% confidence intervals for the dependence model when fitted to the Liver Registration data set. The parameter estimates obtained when fitting the independence model to the Liver Registration data set are included for comparison.

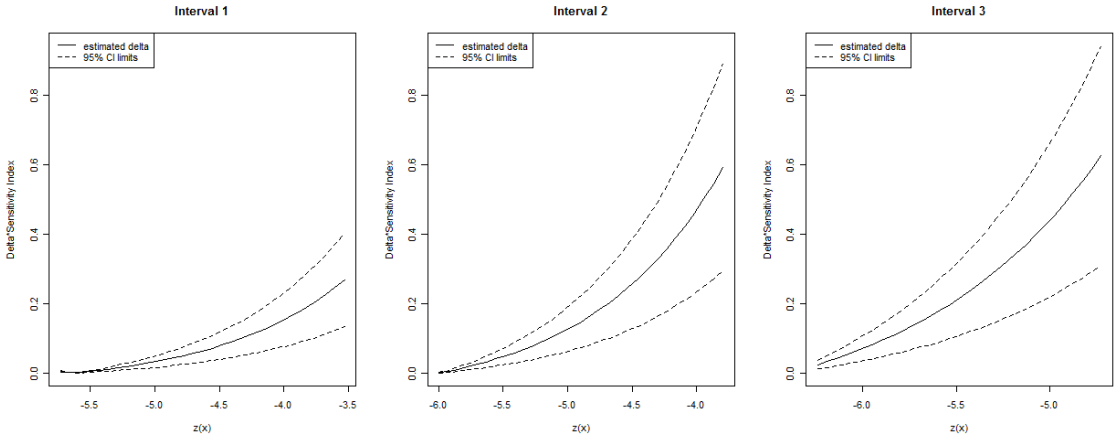

Figure 5.1: The results of the sensitivity analysis for the linear predictor for time to failure using the value of δ estimated by the dependence model

Registration data the discrepancy between the results of the sensitivity analysis and the change in $\hat{w}_\delta(\mathbf{x})$ and $\hat{w}_0(\mathbf{x})$ using the results of the dependent model is small.

The results of the sensitivity analysis for $\boldsymbol{\theta}$ are given in Table 5.4. The estimated values of $\hat{\boldsymbol{\theta}}_{0.2769} - \hat{\boldsymbol{\theta}}_0$ found using the sensitivity analysis are compared to the observed values of $\hat{\boldsymbol{\theta}}_{0.2769} - \hat{\boldsymbol{\theta}}_0$ found by taking the difference of the parameter estimates in Table 5.3. We can see from Table 5.4 that we have mixed results concerning the accuracy of the sensitivity analysis. For most parameters the sensitivity analysis does correctly identify the direction of the change in the parameter estimates. However for patients with metabolic liver disease and white, Asian or black patients this is not the case. Even if the sensitivity analysis correctly identifies the direction of the change, then it may either overestimate or underestimate the magnitude of the change.

Approximate values of $\hat{\boldsymbol{\theta}}_{0.2769}$ can be found by adding the estimated values of $\hat{\boldsymbol{\theta}}_{0.2769} - \hat{\boldsymbol{\theta}}_0$ given in Table 5.4 to the values of $\hat{\boldsymbol{\theta}}_0$ from Table 5.3. These values of $\hat{\boldsymbol{\theta}}_{0.2769}$ can then be used to find the change in the estimated linear predictor for T under this sensitivity analysis. This is done for each individual in the data set using the expression

$$\hat{w}_{0.2769}(\mathbf{x}_{ij}) - \hat{w}_0(\mathbf{x}_{ij}) = \hat{\boldsymbol{\theta}}'_{0.2769}\mathbf{x}_{ij} - \hat{\boldsymbol{\theta}}'_0\mathbf{x}_{ij}.$$

The largest value of this change that is estimated by the sensitivity analysis for $\boldsymbol{\theta}$ is 0.3869. This is very close to the observed change in the estimated linear predictor which was 0.3868. These results suggest that the sensitivity analysis for $\boldsymbol{\theta}$ is more accurate than the sensitivity analysis for $w(\mathbf{x})$. Therefore, the sensitivity analysis for $\boldsymbol{\theta}$ should be used when we wish to apply a sensitivity analysis to a piecewise exponential model with covariates.

Parameter	Estimated values of $\hat{\boldsymbol{\theta}}_{0.2769} - \hat{\boldsymbol{\theta}}_0$	Observed value of $\hat{\boldsymbol{\theta}}_{0.2769} - \hat{\boldsymbol{\theta}}_0$
Intercept	0.6085	0.4908
PLD - PBC	-0.0317	-0.0230
PLD - PSC	-0.0161	-0.0061
PLD - ALD	-0.0110	-0.0061
PLD - AID	-0.0280	-0.0450
PLD - HCV	-0.0300	-0.0354
PLD - HBV	-0.0279	-0.0242
PLD - Cancer	-0.0303	-0.0388
PLD - Metabolic	0.0096	-0.0166
PLD - Other	-0.0192	-0.1048
Ethnicity - White	-0.0647	0.0975
Ethnicity - Asian	-0.0402	0.0871
Ethnicity - Black	-0.0901	0.0551
Ethnicity - Chinese	0.0019	0.2219
UKELD	-0.0036	-0.0055
Age	-0.0005	-0.0004
j - Interval 1	-0.2617	-0.1887
j - Interval 2	-0.2198	-0.1458

Table 5.4: Comparison of the estimated values of $\hat{\boldsymbol{\theta}}_{0.2769} - \hat{\boldsymbol{\theta}}_0$ found using the sensitivity analysis from 4.3.2 with the observed values of $\hat{\boldsymbol{\theta}}_{0.2769} - \hat{\boldsymbol{\theta}}_0$ found using the values in Table 5.3.

5.2 Simulations

To investigate the accuracy of the sensitivity analysis for the difference in the parameters from the independence and dependence models, over a range of parameter values, a simulation study was conducted. The main aim is to establish whether the actual difference in the parameter estimates is generally overestimated or underestimated by the sensitivity analysis. We also wish to assess the accuracy of the sensitivity analysis as a function of dependence. This would give us some idea of when the use of this method is appropriate.

To keep the computation simple, we shall generate data from a piecewise exponential distribution with no other covariates. The sensitivity analysis will be carried out on the simulated data along with fitting the dependent model so the accuracy of our method can be assessed. This will be done for a wide range of the parameters θ, γ and δ, so we can identify the situations where use of the sensitivity analysis is appropriate. The models fitted to the simulated dataset will assume the same distribution that the dataset was simulated from. This allows us to assess the accuracy of the sensitivity analysis when we have fitted the correct model.

5.2.1 Simulation study set-up

The different combinations of these parameters used in the simulations are given in Table 5.5. Each of these combinations was combined with δ values of -0.4, -0.3, -0.2, -0.1, 0.1, 0.2, 0.3 and 0.4. For each different combination of θ, γ and δ we simulated 500 replicates. In all the simulations, we assume $n = 2000$. For each scenario in Table 5.5, we simulate observations from a 2-interval piecewise exponential model. We use e^{θ} as the hazard of failure in the second interval, and $e^{\theta+j_1}$ as the hazard of failure in the first interval. The hazard of censoring in the second interval is e^{γ}, with the hazard in the first interval being $e^{\gamma+j_2}$. An arbitrary cut-point for the 2-interval piecewise exponential model is chosen to give approximately equal numbers of events in the two intervals. These are also given in Table 5.5.

The dependent model will be fitted and the sensitivity analysis will be applied to each simulated data set. When fitting the dependent model, the value of δ will be fixed. This is because there is very little information about δ in the data, even after identifying assumptions have been made, and consistent estimates of δ cannot be obtained. This would make it difficult to make meaningful comparisons between different parameter combinations. Therefore we used a profile likelihood approach when estimating the other parameters in the model. Similarly, the amount of dependence assumed in the sensitivity analysis is the fixed value of δ used when fitting the dependent model.

For each replication the parameter estimates from the dependent model, $\hat{\boldsymbol{\theta}}_{\delta}^{(d)}$, were

Part	θ	j_1	γ	j_2	Cut-point used
1	-2	-1	-3	-0.5	15
2	-4	-1	-3	-0.5	30
3	-6	-1	-3	-0.5	30
4	-2	-1	-4	-0.5	15
5	-4	-1	-4	-0.5	50
6	-6	-1	-4	-0.5	75
7	-2	-1	-5	-0.5	15
8	-4	-1	-5	-0.5	70
9	-6	-1	-5	-0.5	150
10	-8	-1	-5	-0.5	150
11	-2	-1	-6	-0.5	15
12	-4	-1	-6	-0.5	90
13	-6	-1	-6	-0.5	300
14	-8	-1	-6	-0.5	400

Table 5.5: Table showing the combinations of θ, γ and cut points used in the simulation study. For each scenario, δ values of -0.4, -0.3, -0.2, -0.1, 0.1, 0.2, 0.3 and 0.4 will be investigated.

found along with the parameter estimates approximated by the sensitivity analysis, $\hat{\boldsymbol{\theta}}_\delta^{(s)}$. The value $\mathbf{D} = (\hat{\boldsymbol{\theta}}_\delta^{(d)} - \hat{\boldsymbol{\theta}}_0) - (\hat{\boldsymbol{\theta}}_\delta^{(s)} - \hat{\boldsymbol{\theta}}_0)$ is of interest. The element of $\mathbf{D}$ with the largest magnitude is found as this corresponds to the largest discrepancy between the results of the dependent model and the results of the sensitivity analysis. If this term is negative, then the sensitivity analysis overestimates the change in the parameter estimates.

Generating from a piecewise exponential distribution

For the sake of simplicity, we shall consider only the piecewise exponential distribution with 2 intervals in our simulations. Zhou (2001) gives an algorithm to transform standard exponential random variables into piecewise exponential random variables,

$$[\text{Generate } Y \sim \exp(1)] \quad \Rightarrow \quad \begin{cases} \text{if } [Y \leq a_1\lambda_1] \text{ return } Y/\lambda_1 \\ \text{if } [Y > a_1\lambda_1] \text{ return } a_1 + (Y - a_1\lambda_1)/\lambda_2, \end{cases} \tag{5.6}$$

where a_1, is the endpoint of the first intervals and λ_1 and λ_2 are the rates in the first and second intervals respectively. As standard exponential random variables can be easily generated using standard software packages, this algorithm is easy to implement.

For the failure time distribution, we will let the rates in the two intervals be e^{θ_1} and e^{θ_2} respectively and simply apply the above algorithm. Simulating the observations from the conditional distribution for the censored observations is a little tricky, due to the

dependence on the failure time distribution. As defined earlier, the parameter for the conditional distribution is $\gamma + \delta i_\gamma^{-1/2} B(t,\theta)$. Under the structure we are using, $i_\gamma = 1$ and $B(t,\theta) = 1 - H_T(t,\theta)$. Under the piecewise exponential model, we can write the cumulative hazard function as

$$H_T(t,\theta) = \sum_{j=1}^{j(t)} t_j e^{\theta_j}$$

where $j(t)$ is the interval number in which the failure occurs and t_j is the time experienced in the jth interval. This means that the rates used to generate observations from the censoring time distribution are

$$\exp\left\{\gamma_1 + \delta\left(1 - \sum_{j=1}^{j(t)} t_j e^{\theta_j}\right)\right\} \quad \text{and} \quad \exp\left\{\gamma_2 + \delta\left(1 - \sum_{j=1}^{j(t)} t_j e^{\theta_j}\right)\right\}.$$

5.2.2 Results

The mean values of the largest element of **D** were calculated along with a 95% confidence interval for the mean using the set-up described in Section 5.2.1. Table 5.6 gives these results for the simulations. These results are also summarised graphically in Figure 5.2. The plots in Figure 5.2 show the effect of θ on the mean observed as δ increases, at each different level of γ.

The majority of the means observed in Table 5.6 are negative, which means that generally the sensitivity analysis overestimates the change in the parameter estimates. From the plots in Figure 5.2, it can be seen that generally we observe the larger means when δ is greater than 0.3, and γ and θ are similar in size or $\gamma > \theta$. The greater the difference between γ and θ, the bigger the mean difference we observe. As the size of γ relative to θ increases, the hazard rate of censoring is also increasing relative to the hazard rate for failure. So the simulated data sets would generally contain an increasing proportion of censored observations. Therefore we observe the largest changes in the mean of **D** when there is a relatively large proportion of censored observations in the data set. Also as the magnitude of δ increases, the size of the mean also increases, especially in the situations with relatively large amounts of censoring.

Analysis of Variance

To establish the effects of the individual parameters on the simulation results, an analysis of variance model that included all the main effects and interactions between θ, γ and δ was fitted. The ANOVA finds that there is a significant 3 factor interaction between δ, θ and γ, as we can see in Table 5.7. As θ increases the mean observed generally decreases, but the 3 factor interaction means that the values of δ and γ will affect the rate at which

δ	$\theta=-2,\ \gamma=-3$	$\theta=-4,\ \gamma=-3$	$\theta=-6,\ \gamma=-3$
-0.4	-0.1492(-0.1514,-0.1470)	-0.1636(-0.1655,-0.1618)	-0.1410(-0.1428,-0.1393)
-0.3	-0.1006(-0.1021,-0.0991)	-0.1239(-0.1253,-0.1226)	-0.1120(-0.1131,-0.1108)
-0.2	-0.0574(-0.0582,-0.0565)	-0.0807(-0.0816,-0.0798)	-0.0767(-0.0774,-0.0760)
-0.1	-0.0234(-0.0238,-0.0230)	-0.0387(-0.0391,-0.0383)	-0.0385(-0.0388,-0.0382)
0.1	-0.0127(-0.0130,-0.0124)	-0.0320(-0.0323,-0.0316)	-0.0360(-0.0364,-0.0356)
0.2	-0.0162(-0.0167,-0.0157)	-0.0544(-0.0555,-0.0533)	-0.0597(-0.0610,-0.0584)
0.3	-0.0144(-0.0150,-0.0137)	-0.0492(-0.0517,-0.0466)	-0.0627(-0.0676,-0.0577)
0.4	-0.0140(-0.0144,-0.0135)	-0.0009(-0.0048,0.0030)	-0.1469(-0.1841,-0.1096)
δ	$\theta=-2,\ \gamma=-4$	$\theta=-4,\ \gamma=-4$	$\theta=-6,\ \gamma=-4$
-0.4	-0.0647(-0.0657,-0.0638)	-0.1249(-0.1262,-0.1236)	-0.1344(-0.1360,-0.1328)
-0.3	-0.0409(-0.0416,-0.0403)	-0.0891(-0.0900,-0.0881)	-0.1044(-0.1054,-0.1033)
-0.2	-0.0224(-0.0228,-0.0220)	-0.0553(-0.0559,-0.0547)	-0.0714(-0.0721,-0.0708)
-0.1	-0.0086(-0.0088,-0.0084)	-0.0250(-0.0253,-0.0247)	-0.0356(-0.0359,-0.0353)
0.1	-0.0045(-0.0046,-0.0043)	-0.0196(-0.0198,-0.0193)	-0.0329(-0.0333,-0.0326)
0.2	-0.0059(-0.0062,-0.0057)	-0.0346(-0.0351,-0.0340)	-0.0549(-0.0562,-0.0537)
0.3	-0.0059(-0.0062,-0.0056)	-0.0408(-0.0420,-0.0397)	-0.0579(-0.0609,-0.0548)
0.4	-0.0064(-0.0066,-0.0061)	-0.0431(-0.0445,-0.0418)	-0.0863(-0.0957,-0.0770)
δ	$\theta=-2,\ \gamma=-5$	$\theta=-4,\ \gamma=-5$	$\theta=-6,\ \gamma=-5$
-0.4	-0.0278(-0.0283,-0.0273)	-0.0806(-0.0815,-0.0798)	-0.1093(-0.1107,-0.1080)
-0.3	-0.0163(-0.0166,-0.0159)	-0.0537(-0.0543,-0.0531)	-0.0836(-0.0844,-0.0827)
-0.2	-0.0084(-0.0086,-0.0082)	-0.0316(-0.0320,-0.0312)	-0.0562(-0.0567,-0.0556)
-0.1	-0.0032(-0.0033,-0.0031)	-0.0134(-0.0136,-0.0132)	-0.0276(-0.0279,-0.0274)
0.1	-0.0017(-0.0018,-0.0016)	-0.0094(-0.0095,-0.0092)	-0.0250(-0.0253,-0.0246)
0.2	-0.0023(-0.0024,-0.0022)	-0.0152(-0.0155,-0.0149)	-0.0415(-0.0426,-0.0405)
0.3	-0.0027(-0.0029,-0.0025)	-0.0183(-0.0187,-0.0178)	-0.0506(-0.0526,-0.0487)
0.4	-0.0030(-0.0032,-0.0029)	-0.0204(-0.0208,-0.0200)	-0.0606(-0.0641,-0.0571)
δ	$\theta=-8,\ \gamma=-5$	$\theta=-2,\ \gamma=-6$	$\theta=-4,\ \gamma=-6$
-0.4	-0.0913(-0.0928,-0.0898)	-0.0108(-0.0111,-0.0105)	-0.0477(-0.0483,-0.0470)
-0.3	-0.0735(-0.0747,-0.0724)	-0.0060(-0.0062,-0.0058)	-0.0301(-0.0306,-0.0297)
-0.2	-0.0502(-0.0510,-0.0494)	-0.0030(-0.0031,-0.0029)	-0.0167(-0.0170,-0.0164)
-0.1	-0.0262(-0.0267,-0.0258)	-0.0011(-0.0012,-0.0011)	-0.0067(-0.0069,-0.0066)
0.1	-0.0226(-0.0231,-0.0221)	-0.00060(-0.00065,-0.00056)	-0.0040(-0.0041,-0.0038)
0.2	-0.0339(-0.0350,-0.0328)	-0.00096(-0.00103,-0.00089)	-0.0057(-0.0059,-0.0055)
0.3	-0.0579(-0.0655,-0.0502)	-0.0012(-0.0013,-0.0011)	-0.0062(-0.0065,-0.0060)
0.4	-0.2225(-0.2677,-0.1773)	-0.0014(-0.0015,-0.0013)	-0.0069(-0.0072,-0.0067)
δ	$\theta=-6,\ \gamma=-6$	$\theta=-8,\ \gamma=-6$	
-0.4	-0.0966(-0.0976,-0.0956)	-0.0907(-0.0923,-0.0891)	
-0.3	-0.0704(-0.0711,-0.0697)	-0.0732(-0.0746,-0.0719)	
-0.2	-0.0442(-0.0446,-0.0438)	-0.0518(-0.0526,-0.0509)	
-0.1	-0.0203(-0.0205,-0.0201)	-0.0266(-0.0270,-0.0262)	
0.1	-0.0174(-0.0177,-0.0172)	-0.0227(-0.0232,-0.0223)	
0.2	-0.0329(-0.0336,-0.0323)	-0.0358(-0.0369,-0.0347)	
0.3	-0.0426(-0.0437,-0.0415)	-0.0567(-0.0613,-0.0521)	
0.4	-0.0508(-0.0521,-0.0494)	-0.1441(-0.1558,-0.1325)	

Table 5.6: The mean of largest element of $\mathbf{D}$ (with 95% confidence intervals) for each combination of parameters given in Table 5.5 and each value of δ.

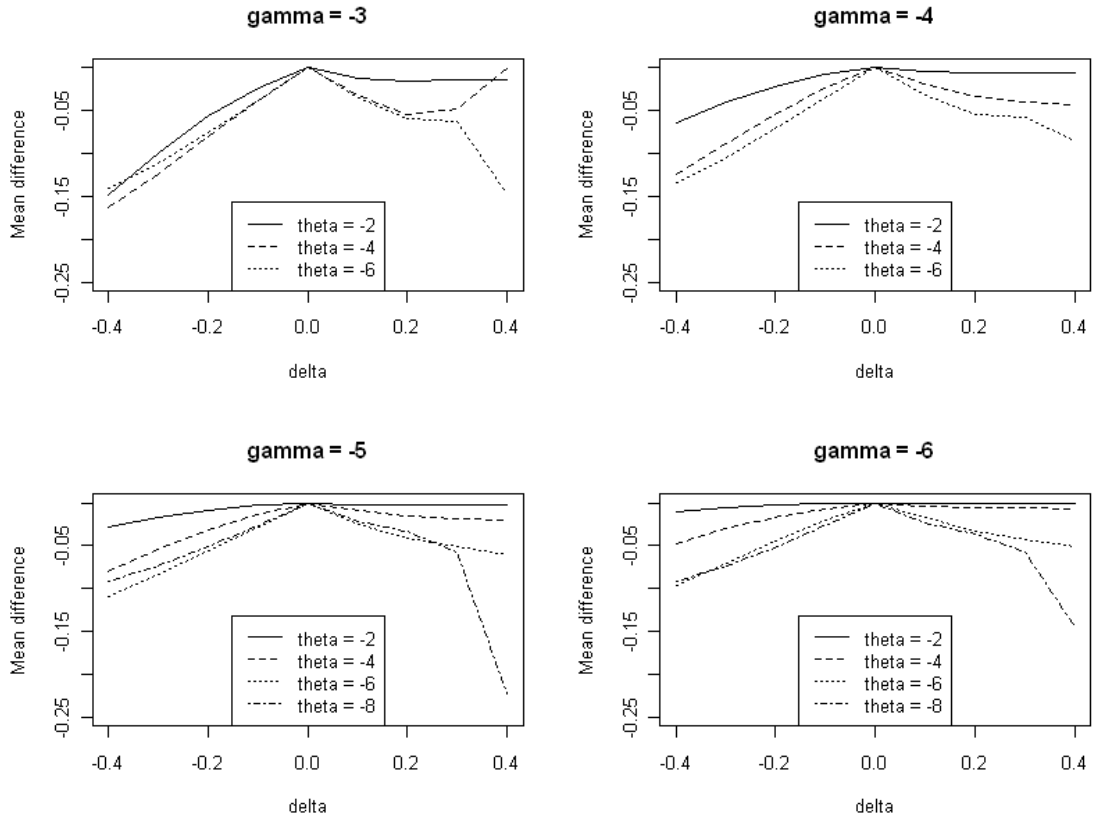

Figure 5.2: Effect of θ on mean of largest element of $\mathbf{D}$ as δ varies between -0.4 and 0.4 for the values of γ considered in the simulation study.

the mean decreases. There is a greater rate of decrease for larger values of γ when δ is negative. This is because as θ increases there is a greater decrease in the proportion of censored observations in the simulated data sets with larger values of γ. Conversely, for positive δ, the decrease in the proportion of censored observations as θ increases is greatest for the simulated data sets with smaller values of γ, so these values have the greatest rate of decrease in the mean. The rate of decrease in the mean is also affected by the magnitude of δ. For larger magnitudes of δ, the rate of decrease in the mean as θ increases is larger. This makes intuitive sense as the effect of censored observations on the results of the sensitivity analysis increases as the magnitude of δ increases.

Parameter	DF	Type I SS	Mean Square	F Value	p
delta	8	48.4913	6.0614	1523.82	< 0.0001
theta	3	23.2965	7.7655	1952.22	< 0.0001
gamma	3	7.7221	2.5740	647.10	< 0.0001
delta*theta	24	19.2382	0.8016	201.52	< 0.0001
delta*gamma	24	8.7563	0.3648	91.72	< 0.0001
theta*gamma	7	0.6545	0.0935	23.51	< 0.0001
delta*theta*gamma	56	5.7415	0.1025	25.77	< 0.0001

Table 5.7: Significance levels of parameters and interactions in analysis of variance

There are a handful of situations that are found to have means that are significantly different from most of the other means. These are when $\theta = -6$ and $\gamma = -3$, $\theta = -8$ and $\gamma = -5$ and $\theta = -8$ and $\gamma = -6$, all for $\delta = 0.4$. It is easy to identify these cases in Figure 5.2.

Further investigations revealed that in these cases, some of the data sets had large outlying values that caused a large increase in the value of the sensitivity index, U. This meant that the sensitivity analysis performed particularly badly for these data sets, resulting in an increased mean for $\mathbf{D}$. This tells us that the accuracy of the sensitivity analysis is affected by the size of the observations included in each interval. This was observed in Sections 4.3.1 and 4.3.2 as the widest interval had the largest estimated changes in the parameter estimates both when applying the sensitivity analysis for scalar parameters and when including covariates. So the accuracy of the sensitivity analysis for piecewise exponential models can be improved by dividing the time into a larger number of small intervals.

The results of the simulation study carried out in this section suggest that the sensitivity analysis is not a good approximation of the change in parameter estimates when there is heavy censoring and δ becomes large. This could help to explain why the sensitivity analysis overestimated the actual change in parameter estimates for the Liver Registra-

tion dataset. In this dataset there is heavy censoring, with 71.7% of patients having a potentially informatively censored time and a further 12.4% having a non-informatively censored time. Also the dependent model fitted suggests that δ is around 0.3, although with a wide confidence interval because even after our assumptions to identify the joint distribution of T and C we have little information about the dependence parameter.

However, even though some situations have been identified where the sensitivity analysis does not give a good approximation to the dependent model, the simulation study in this section shows that there are many situations when the sensitivity analysis does provide a reasonable approximation to the dependent model. This means that while the sensitivity analysis was not as accurate as we would have hoped for the Liver Registration data, it is still suitable for application in other situations.

5.3 Inclusion of Extra Terms in Approximations used in Sensitivity Analysis

There are several Taylor expansions that are used in the derivation of the sensitivity analysis described in Chapter 4. The accuracy of the sensitivity analysis may be improved by including extra terms in any of these expansions. However, it is still necessary that there is a closed form equation for the difference in the parameter estimates. Because of this restriction we found that it is possible to include extra terms in the approximation in (4.3), but not those in used in (4.7).

Here, an equation for the sensitivity analysis is derived when using an additional quadratic term in the approximation for the conditional density function of C. Hence (4.3) is replaced by

$$\begin{aligned} f_{T,C}(t_j,c_j) \simeq & f_T(t_j,\theta_j)\Bigg[f_C(c_j,\gamma_j)+\delta i_{\gamma_j}^{-1/2}B(t_j,\theta_j)\frac{\partial}{\partial\gamma_j}f_C(c_j,\gamma_j) \\ & \qquad +\frac{1}{2}(\delta i_{\gamma_j}^{-1/2}B(t_j,\theta_j))^2\frac{\partial^2}{\partial\gamma^2}f_C(c_j,\gamma_j)\Bigg] \\ = & f_T(t_j,\theta_j)f_C(c_j,\gamma_j)\Bigg[1+\delta i_{\gamma_j}^{-1/2}B(t_j,\theta_j)\frac{\frac{\partial}{\partial\gamma_j}f_C(c_j,\gamma_j)}{f_C(c_j,\gamma_j)} \\ & \qquad +\frac{1}{2}\delta^2 i_{\gamma_j}^{-1}B(t_j,\theta_j)^2\frac{\frac{\partial^2}{\partial\gamma_j^2}f_C(c_j,\gamma_j)}{f_C(c_j,\gamma_j)}\Bigg]. \end{aligned}$$

If this approximation of the joint density function is used in the likelihood in (4.1), then

the likelihood becomes

$$\begin{aligned}
\ell_\delta(\boldsymbol{\theta},\boldsymbol{\gamma}) &\simeq \ell_0(\boldsymbol{\theta},\boldsymbol{\gamma}) \\
&+ \delta \sum_{i=1}^{n}\sum_{j=1}^{m} i_\gamma^{-1/2}\Bigg\{ I_{ij}\left[B(y_{ij},\theta_j)\frac{\frac{\partial}{\partial\gamma_j}S_C(y_{ij},\gamma_j)}{S_C(y_{ij},\gamma_j)} + \frac{1}{2}\delta i_\gamma^{-1/2}B(y_{ij},\theta_j)^2\frac{\frac{\partial^2}{\partial\gamma_j^2}S_C(y_{ij},\gamma_j)}{S_C(y_{ij},\gamma_j)}\right] \\
&\quad + (1-I_{ij})(1-Z_{ij})\left[\mu(y_{ij},\theta_j)\frac{\frac{\partial}{\partial\gamma_j}S_C(y_{ij},\gamma_j)}{S_C(y_{ij},\gamma_j)} + \frac{1}{2}\delta i_\gamma^{-1/2}\nu(y_{ij},\theta_j)\frac{\frac{\partial^2}{\partial\gamma_j^2}S_C(y_{ij},\gamma_j)}{S_C(y_{ij},\gamma_j)}\right] \\
&\quad + Z_{ij}(1-I_{ij})\left[\mu(y_{ij},\theta_j)\frac{\frac{\partial}{\partial\gamma_j}f_C(y_{ij},\gamma_j)}{f_C(y_{ij},\gamma_j)} + \frac{1}{2}\delta i_\gamma^{-1/2}\nu(y_{ij},\theta_j)\frac{\frac{\partial^2}{\partial\gamma_j^2}f_C(y_{ij},\gamma_j)}{f_C(y_{ij},\gamma_j)}\right]\Bigg\}, \qquad (5.7)
\end{aligned}$$

where

$$\begin{aligned}
\mu(y_{ij},\theta_j) &= \frac{\int_{y_{ij}}^{\infty} B(u,\theta_j)f_T(u,\theta_j)du}{S_T(y_{ij},\theta_j)}, \quad \text{and} \\
\nu(y_{ij},\theta_j) &= \frac{\int_{y_{ij}}^{\infty} B(u,\theta_j)^2 f_T(u,\theta_j)du}{S_T(y_{ij},\theta_j)}.
\end{aligned}$$

The method used to derive the sensitivity analysis equation is similar to that for the sensitivity analysis for $\boldsymbol{\theta}$ in Section 4.2.3. This means that covariates will be included in the sensitivity analysis. So θ and γ will be replaced by the parameter vectors $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}$ respectively. However the linear predictor $\eta = \boldsymbol{\gamma}'\mathbf{x}$ To obtain an expression for $\hat{\boldsymbol{\theta}}_\delta - \hat{\boldsymbol{\theta}}_0$, Taylor expansions of the vector score functions

$$\boldsymbol{r}_0(\boldsymbol{\theta}_0) = \frac{\partial}{\partial\boldsymbol{\theta}}\ell_0(\boldsymbol{\theta},\eta,\mathbf{x}) \quad \text{and} \quad \boldsymbol{r}_\delta(\hat{\boldsymbol{\theta}}_\delta) = \frac{\partial}{\partial\boldsymbol{\theta}}\ell_\delta(\boldsymbol{\theta},\eta,\mathbf{x}).$$

need to be used. As in Chapter 4, these expansions will only include the linear terms, so that

$$\begin{aligned}
\boldsymbol{r}_0(\hat{\boldsymbol{\theta}}_0) &\simeq \boldsymbol{r}_0(\boldsymbol{\theta}) - (\hat{\boldsymbol{\theta}}_0 - \boldsymbol{\theta})i(\boldsymbol{\theta},\mathbf{x}) = 0 \\
\boldsymbol{r}_\delta(\hat{\boldsymbol{\theta}}_\delta) &\simeq \boldsymbol{r}_\delta(\boldsymbol{\theta}) - (\hat{\boldsymbol{\theta}}_\delta - \boldsymbol{\theta})i(\boldsymbol{\theta},\mathbf{x}) = 0, \qquad (5.8)
\end{aligned}$$

where the (k,l)th element of the information matrix $i(\boldsymbol{\theta},\mathbf{x})$ is

$$\frac{\partial^2}{\partial\theta_k\partial\theta_l}\ell_0(\boldsymbol{\theta},\eta,\mathbf{x}).$$

The expressions in (5.8) can be rearranged to give

$$\hat{\boldsymbol{\theta}}_\delta - \hat{\boldsymbol{\theta}}_0 \simeq i(\boldsymbol{\theta},\mathbf{x})^{-1}(\boldsymbol{r}_\delta(\boldsymbol{\theta}) - \boldsymbol{r}_0(\boldsymbol{\theta})), \qquad (5.9)$$

using the likelihood in (5.7) to obtain the kth component of $\boldsymbol{r}_\delta(\boldsymbol{\theta}) - \boldsymbol{r}_0(\boldsymbol{\theta})$, which is

$$\begin{aligned}
\delta \sum_{i=1}^{n}\sum_{j=1}^{m} i_\eta^{-\frac{1}{2}} \Bigg\{ & I_{ij}\Bigg[\frac{\partial B(y_{ij},\boldsymbol{\theta},\mathbf{x}_{ij})}{\partial\theta_k}\frac{\frac{\partial}{\partial\eta_{ij}}S_C(y_{ij},\eta_{ij})}{S_C(y_{ij},\eta_{ij})} \\
& \qquad + \frac{1}{2}\delta i_\eta^{-\frac{1}{2}}\frac{\partial B(y_{ij},\boldsymbol{\theta},\mathbf{x}_{ij})^2}{\partial\theta_k}\frac{\frac{\partial^2}{\partial\eta_{ij}^2}S_C(y_{ij},\eta_{ij})}{S_C(y_{ij},\eta_{ij})}\Bigg] \\
& + (1-I_{ij})(1-Z_{ij})\Bigg[\frac{\partial \mu(y_{ij},\boldsymbol{\theta},\mathbf{x}_{ij})}{\partial\theta_k}\frac{\frac{\partial}{\partial\eta_{ij}}S_C(y_{ij},\eta_{ij})}{S_C(y_{ij},\eta_{ij})} \\
& \qquad + \frac{1}{2}\delta i_\eta^{-\frac{1}{2}}\frac{\partial \nu(y_{ij},\boldsymbol{\theta},\mathbf{x}_{ij})}{\partial\theta_k}\frac{\frac{\partial^2}{\partial\eta_{ij}^2}S_C(y_{ij},\eta_{ij})}{S_C(y_{ij},\eta_{ij})}\Bigg] \\
& + Z_{ij}(1-I_{ij})\Bigg[\frac{\partial \mu(y_{ij},\boldsymbol{\theta},\mathbf{x}_{ij})}{\partial\theta_k}\frac{\frac{\partial}{\partial\eta_{ij}}f_C(y_{ij},\eta_{ij})}{f_C(y_{ij},\eta_{ij})} \\
& \qquad + \frac{1}{2}\delta i_\eta^{-\frac{1}{2}}\frac{\partial \nu(y_{ij},\boldsymbol{\theta},\mathbf{x}_{ij})}{\partial\theta_k}\frac{\frac{\partial^2}{\partial\eta_{ij}^2}f_C(y_{ij},\eta_{ij})}{f_C(y_{ij},\eta_{ij})}\Bigg]\Bigg\}. \qquad (5.10)
\end{aligned}$$

It is possible to apply the proportional hazards structure that was outlined in Section 4.2.2. As before, if this structure is assumed, $i_\eta = 1$, $B(t_j,\boldsymbol{\theta},\mathbf{x}_j) = 1 - H_T(t_j,\boldsymbol{\theta},\mathbf{x}_j)$ and $\mu(t_j,\boldsymbol{\theta},\mathbf{x}_j) = -H_T(t_j,\boldsymbol{\theta},\mathbf{x}_j)$. The form of $\nu(t_j,\boldsymbol{\theta},\mathbf{x}_j)$ is $1 + H_T(t_j,\boldsymbol{\theta},\mathbf{x}_j)^2$. Using $\frac{\partial}{\partial\theta_k}H_T(t_j,\boldsymbol{\theta},\mathbf{x}_j) = x_k H_T(t_j,\boldsymbol{\theta},\mathbf{x}_j)$ and $\frac{\partial}{\partial\eta}H_C(c_j,\eta_j) = H_C(c_j,\eta_j)$, then the derivatives of $B(t_j,\boldsymbol{\theta},\mathbf{x}_j)$, $B(t_j,\boldsymbol{\theta},\mathbf{x}_j)^2$, $\mu(t_j,\boldsymbol{\theta},\mathbf{x}_j)$ and $\nu(t_j,\boldsymbol{\theta},\mathbf{x}_j)$ are

$$\begin{aligned}
\frac{\partial}{\partial\theta_k}B(t_j,\boldsymbol{\theta},\mathbf{x}_j) &= -x_k H_T(t_j,\boldsymbol{\theta},\mathbf{x}_j), \\
\frac{\partial}{\partial\theta_k}B(t_j,\boldsymbol{\theta},\mathbf{x}_j)^2 &= -2x_k H_T(t_j,\boldsymbol{\theta},\mathbf{x}_j) + 2x_k H_T(t_j,\boldsymbol{\theta},\mathbf{x}_j)^2, \\
\frac{\partial}{\partial\theta_k}\mu(t_j,\boldsymbol{\theta},\mathbf{x}_j) &= -x_k H_T(t_j,\boldsymbol{\theta},\mathbf{x}_j) \quad \text{and} \\
\frac{\partial}{\partial\theta_k}\nu(t_j,\boldsymbol{\theta},\mathbf{x}_j) &= 2x_k H_T(t_j,\boldsymbol{\theta},\mathbf{x}_j)^2.
\end{aligned}$$

It can also be shown that

$$\begin{aligned}
\frac{\frac{\partial}{\partial z(\mathbf{x}_j)}S_C(c_j,z(\mathbf{x}_j))}{S_C(c_j,z(\mathbf{x}_j))} &= -H_C(c_j,z(\mathbf{x}_j)), \\
\frac{\frac{\partial^2}{\partial z(\mathbf{x}_j)^2}S_C(c_j,z(\mathbf{x}_j))}{S_C(c_j,z(\mathbf{x}_j))} &= H_C(c_j,z(\mathbf{x}_j))(H_C(c_j,z(\mathbf{x}_j)) - 1), \\
\frac{\frac{\partial}{\partial z(\mathbf{x}_j)}f_C(c_j,z(\mathbf{x}_j))}{f_C(c_j,z(\mathbf{x}_j))} &= 1 - H_C(c_j,z(\mathbf{x}_j)) \quad \text{and} \\
\frac{\frac{\partial^2}{\partial z(\mathbf{x}_j)^2}f_C(c_j,z(\mathbf{x}_j))}{f_C(c_j,z(\mathbf{x}_j))} &= 1 - 3H_C(c_j,z(\mathbf{x}_j)) + H_C(c_j,z(\mathbf{x}_j))^2.
\end{aligned}$$

If these terms are substituted into (5.10), then a simplified version of the kth component

of $\boldsymbol{r}_\delta(\boldsymbol{\theta}) - \boldsymbol{r}_0(\boldsymbol{\theta})$ is obtained, which is

$$\begin{aligned}
&\delta \sum_{i=1}^{n}\sum_{j=1}^{m} x_{ijk}\Big\{H_T(y_{ij},\boldsymbol{\theta},\mathbf{x}_{ij})H_C(y_{ij},\eta_{ij}) - Z_{ij}(1-I_{ij})H_T(y_{ij},\boldsymbol{\theta},\mathbf{x}_{ij})\Big\} \\
+\delta^2 &\sum_{i=1}^{n}\sum_{j=1}^{m} x_{ijk}\Big\{(H_T(y_{ij},\boldsymbol{\theta},\mathbf{x}_{ij}) - I_{ij})H_T(y_{ij},\boldsymbol{\theta},\mathbf{x}_{ij})H_C(y_{ij},\eta_{ij})(H_C(y_{ij},\eta_{ij})-1) \\
&\qquad + Z_{ij}(1-I_{ij})H_T(y_{ij},\boldsymbol{\theta},\mathbf{x}_{ij})^2(1-2H_C(y_{ij},\eta_{ij}))\Big\}. \qquad (5.11)
\end{aligned}$$

5.3.1 Application of sensitivity analysis that uses additional terms in approximations to the Liver Registration data set

The sensitivity analysis using (5.9) is applied to the Liver Registration data set to see if it gives an improvement on the sensitivity analysis presented in Chapter 4. Again, it is assumed that the lifetime and censoring variables each have piecewise exponential marginal distributions, each with three intervals with cut points at 40 and 165 days. Age, recipient ethnicity, primary liver disease category and UKELD score at time of registration were also included as covariates in the model for time to death. The model for time to censoring only included an intercept term so that the results of this sensitivity analysis can be compared to the observed values from the fitted dependence model.

The hazards and associated functions for T and C with piecewise exponential marginal distributions can be expressed as:

$$\begin{aligned}
h_T(t_j,\boldsymbol{\theta},\mathbf{x}_j) &= e^{\boldsymbol{\theta}'\mathbf{x}_j} & h_C(c_j,\eta_j) &= e^{\eta_j} \\
H_T(t_j,\boldsymbol{\theta},\mathbf{x}_j) &= e^{\boldsymbol{\theta}'\mathbf{x}_j}t_j & H_C(c_j,\eta_j) &= e^{\eta_j}c_j \\
S_T(t_j,\boldsymbol{\theta},\mathbf{x}_j) &= \exp(-e^{\boldsymbol{\theta}'\mathbf{x}_j}t_j) & S_C(c_j,\eta_j) &= \exp(-e^{\eta_j}c_j) \qquad (5.12)
\end{aligned}$$

If the forms in (5.12) are substituted in (5.11), then the final form of the kth component of $\boldsymbol{r}_\delta(\boldsymbol{\theta}) - \boldsymbol{r}_0(\boldsymbol{\theta})$ that shall be used in (5.9) is

$$\begin{aligned}
&\delta \sum_{i=1}^{n}\sum_{j=1}^{m} x_{ijk}e^{\boldsymbol{\theta}'\mathbf{x}_{ij}}\left\{e^{\eta_{ij}}y_{ij}^2 - Z_{ij}(1-I_{ij})y_{ij}\right\} \\
&\quad + \delta^2\Bigg\{\sum_{i=1}^{n}\sum_{j=1}^{m} x_{ijk}e^{\boldsymbol{\theta}'\mathbf{x}_{ij}}\left\{(e^{\boldsymbol{\theta}'\mathbf{x}_{ij}}y_{ij} - I_{ij})e^{\eta_{ij}}y_{ij}^2(e^{\eta_{ij}}y_{ij}-1)\right. \\
&\qquad\qquad \left. + Z_{ij}(1-I_{ij})e^{\boldsymbol{\theta}'\mathbf{x}_{ij}}y_{ij}^2(1-2e^{\eta_{ij}}y_{ij})\right\}\Bigg\}, \qquad (5.13)
\end{aligned}$$

with the (k,l)th element of the information matrix $i(\boldsymbol{\theta},\mathbf{x})$ becoming

$$\sum_{i=1}^{n}\sum_{j=1}^{m} x_{ijk}x_{ijl}e^{\boldsymbol{\theta}'\mathbf{x}_{ij}}y_{ij}.$$

The first term in (5.13) is the same as (4.30) in Section 4.2.3. So this sensitivity analysis can be viewed as the original piecewise sensitivity analysis with a correction term. To see if having this correction term in the sensitivity analysis improves its accuracy, the results of the sensitivity analysis in Table 5.8 are compared to the values in Table 5.4.

Parameter	$\hat{\boldsymbol{\theta}}_{0.2769} - \hat{\boldsymbol{\theta}}_0$
Intercept	0.6953
PLD - PBC	-0.0757
PLD - PSC	-0.0526
PLD - ALD	-0.0489
PLD - AID	-0.0709
PLD - HCV	-0.0597
PLD - HBV	-0.0805
PLD - Cancer	-0.0641
PLD - Metabolic	0.0415
PLD - Other	-0.0597
Ethnicity - White	-0.0623
Ethnicity - Asian	-0.0465
Ethnicity - Black	-0.0944
Ethnicity - Chinese	0.0035
UKELD	-0.0033
Age	-0.0008
j - Interval 1	-0.3241
j - Interval 2	-0.2793

Table 5.8: The results of the sensitivity analysis that includes extra terms in the approximations used, for $\delta = 0.2769$.

The piecewise sensitivity analysis that uses extra terms overestimates more of the changes in the parameter estimates than the original sensitivity analysis. Also, any values that were already overestimated by the original piecewise sensitivity analysis are overestimated even more by the piecewise sensitivity analysis that uses extra terms, particularly the values corresponding to the intercepts in each interval. Therefore, the sensitivity analysis that uses extra terms is not an improvement on the original piecewise sensitivity analysis.

5.4 Summary

The aim of this chapter is to assess how accurate the sensitivity analysis is overall and to identify any situations where it performs particularly badly.

We detail how to fit the model that includes dependence before any simplifying assumptions. Although it is possible to fit this model, it is not simple and can be very time consuming, especially if there are a large number of parameters. Also the dependence assumption used can not be checked. This highlights why we need the sensitivity analysis as we do not wish to fit these complex models if it is not necessary. When this model was fitted to the Liver Registration data set, we found that the sensitivity analysis for $w(\mathbf{x})$ overestimated the change in parameter estimates. However, there were mixed results for the sensitivity analysis for $\boldsymbol{\theta}$, although it did overestimate the change in the parameter estimates corresponding to the intercepts in each interval. Overall the sensitivity analysis for $\boldsymbol{\theta}$ was found to be more accurate than the sensitivity analysis for $w(\mathbf{x})$.

To assess the general accuracy of the piecewise sensitivity analysis, a simulation study was carried out across a range of parameter combinations, that correspond to a variety of different situations. However, for simplicity, these simulations only consider models with intercepts in each interval for both time to death and time to censoring. The sensitivity analysis does tend to overestimate the change in these parameter estimates, although it is worst when there are large outlying observations in the data set. The sensitivity analysis also tends to overestimate the difference in the parameter estimates corresponding to the intercepts in each interval when a data set has a large amount of censoring and the correlation coefficient between T and C is assumed to be greater than 0.3.

A sensitivity analysis that uses an extra quadratic term in one of its Taylor expansions was derived, as it was hoped this might be able to correct the overestimation seen when there is heavy censoring in the data set. However, it was found for the Liver Registration data set that this sensitivity analysis overestimated the change in the parameter estimates even more than the original piecewise sensitivity analysis.

Chapter 6

Comparing Waiting List and Post-transplant Mortality in Presence of Informative Censoring

We have reviewed the current methods for accounting for informative censoring in Chapters 2 and 3. We have also established that there is a change in parameter estimates in our data set of interest when we assume informative censoring instead of non-informative censoring that is large enough to be of concern. This was done using the sensitivity analysis methodology developed in Chapters 4 and 5. So we will now consider a subject of interest to NHSBT, which is whether patients are expected to receive a benefit from transplantation at all values of the UKELD score. We present a method that answers this question by making use of one of the methods previously considered and apply it to the Liver Registration data set. We also describe how this method can be extended to assess whether patients receive a benefit from alternative therapy transplants, such as using a split liver or a liver from an extended criteria donor. However, this is not applied to the Liver Registration data set.

It is important to be able to show that patients generally have an improvement in their expected survival after transplantation. This becomes especially important when we are considering some of the policies that have been adopted to increase the number of donor livers available such as split livers or extended criteria donor livers. This can be assessed using a concept known as survival benefit, which uses the covariate-adjusted hazard ratio for transplantation compared to not receiving a transplant, to quantify the expected change in post-transplant mortality relative to waiting list mortality. If this ratio is less than 1 then the expected survival of a patient after a transplant is greater then their expected survival if they were to remain on the waiting list.

Here we shall calculate the survival benefit for different groups of UKELD scores to see which patients get the most survival benefit, or if there are any patients that do not experience a significant difference between waiting and post-transplant mortality. This shall be done using a method known as sequential stratification, which sets up experiments to compare the survival of each transplanted patient with similar candidates who were on the waiting list for at least the same amount of time as the transplanted patient. It also uses weights similar to the inverse probability of censoring weights to account for informative censoring.

Firstly, we will introduce the notation required for this method and discuss covariates and the models necessary for the UKELD score before describing the method in more detail. We will then describe the weight function that needs to be used to account for the informative censoring in the data set. An estimating equation for this method is then derived. Finally, we apply the method to our data set to produce estimates of the survival benefit for individuals in different UKELD groups.

6.1 Notation and Covariates

There are many different events that could be observed for each individual transplant patient. Those include D_i, time of death and C_i, time to censoring due to end of study or lost to follow up. T_i will be used define time to transplantation. The observed endpoint for all individuals will be $Y_i = \min(D_i, C_i)$. Therefore a death indicator, $\Delta_i = I(D_i < C_i)$, will be necessary along with an at-risk indicator, $R_i(t) = I(Y_i \geq t)$. Ideally any patients removed from the waiting list will be followed up after this removal. This was possible in Schaubel et al.(2009b) using additional information from the Social Security Death Master File. It is not possible for us to do something similar, as we have incomplete death information for patients who are removed. However we do know whether they were removed due to deteriorating condition or for other reasons. So if an individual was removed because their condition had deteriorated, we assumed that they died on the date of removal. Individuals that were removed for other reasons were censored at the time of removal.

The counting process format will be used in this approach. This means that counting processes for both death and transplantation are set up. These being $N_i^D(t) = I(D_i \leq t, \Delta_i = 1)$ and $N_i^T(t) = I(T_i \leq \min(t, Y_i))$ respectively. These remain at zero until the patient either dies or receives a transplant, at which point they jump to one. The increments of these processes are given by $dN_i^D(t) = N_i^D(t^- + dt) - N_i^D(t^-)$ and $dN_i^T(t) = N_i^T(t^- + dt) - N_i^T(t^-)$ respectively.

Counting processes for the deaths and transplantations in the entire sample can also

be defined as $N^D(t) = \sum_{i=1}^n N_i^D(t)$ and $N^T(t) = \sum_{i=1}^n N_i^T(t)$. These count the number of deaths or transplantations in the sample at or before time t. The increments in these processes can also be found in a similar way to those outlined above.

There will also be the vector $\mathbf{V}(t)$ that contains the values of all the covariates identified as being a significant predictor of time to death for patients on the transplant waiting list at time t. One of the major components of this vector will be the UKELD score. As this score will change over time then will we need the recorded history, $\bar{\mathbf{V}}(t)$, of the vector $\mathbf{V}(t)$.

6.2 Models for UKELD score

Ideally, we have measurements of the UKELD at several different times whilst the patient remains on the waiting list. Unfortunately in the data set we have the UKELD score at registration for all the patients that will be included in this analysis, and only a second reading at time of transplantation for those who receive a transplant. So we have to choose a method of modelling the UKELD scores at interim time points.

For the patients where we have two data points, we consider using linear interpolation to compute the value of the UKELD score at intermediate time points. However, this is probably not the best method to use as UKELD scores do not tend to vary linearly with time. They tend to stay roughly constant until there is a fairly sudden deterioration in the condition of the patient. But we do not have enough UKELD values to capture this behaviour, so the two choices we have are linear interpolation or carrying forward the UKELD values recorded at the time of registration on the waiting list.

Some of the patients with recorded UKELD values at the time of transplant have UKELD values at time of transplant that are lower than their recorded values at time of registration. If we use linear interpolation for these patients, then we would have decreasing values of UKELD at the intermediate times. This would suggest that the patient's liver function is improving over time, which would make it unlikely that they would be given a liver transplant. There is obviously something happening to these patients that the recorded values of UKELD we have for them does not capture, therefore it would be better to assume that their UKELD values remain constant at the value recorded at time of registration.

UKELD Model 1 For patients with UKELD values at time of transplant that are larger then those at time of registration, we use linear interpolation to find intermediate values. For patients who are not transplanted, or have a value of UKELD at transplant that is lower than that at time of registration, we assume that the UKELD value remains

constant at the value recorded at time of registration.

UKELD Model 2 For this model, we assume that the UKELD scores remain constant at the value recorded at time of registration for all patients.

More information on how a patient's UKELD score changes over time is now being collected, so in future applications of this method, these values can be used instead of either of the models we have presented here.

6.3 Sequential Stratification Method

The sequential stratification method allows the comparison of waiting list mortality and post-transplant mortality for liver transplant patients, from which the survival benefit of transplantation can be derived. There are n_T ordered times to transplantation from registration on the list, t_j where $j = 1 \dots n_T$. At each of these transplant times an "experiment" is initiated that compares the mortality of the patient being transplanted with the mortality of similar patients who were still on the waiting list at time t_j. In Schaubel et al.(2009b), similar patients were considered to be those in the same geographic region and the same MELD category, who are still at risk. For the NHSBT data that will be analysed, geographic region does not have the same effect on survival as in the United States so there will be no need to condition on this. Also the UKELD score will be used instead of the MELD score.

Schaubel et al. (2009b) classify patients who are still at risk as those who are alive, untransplanted and active on the waiting list at time t_j. This means that $\min(T_i, Y_i) > t_j$. So, the patient's time to transplantation or some other endpoint from time of registration is greater than t_j. This is what we shall refer to as patients being matched by time from registration. We also consider the case where patients are matched by date of transplant, where patients are only considered to be at risk if they are alive and untransplanted and active on the list on the date of transplant j. Figure 6.1 illustrates some of the differences between these two methods of deciding whether a patient should be included in control group. The dashed arrows in these plots show how long the patients have been on the waiting list before the time of the jth transplant.

When matching by time from registration, each individual has spent the same amount of time on the waiting list prior to being entered into the experiment. This means time from registration can be used in the models that we fit. It is not as simple as that when matching by date of transplant. As we can see from the right hand plot in Figure 6.1, individuals will have been active on the waiting list for different lengths of time. This means that to be able to make meaningful comparisons between the individuals in an

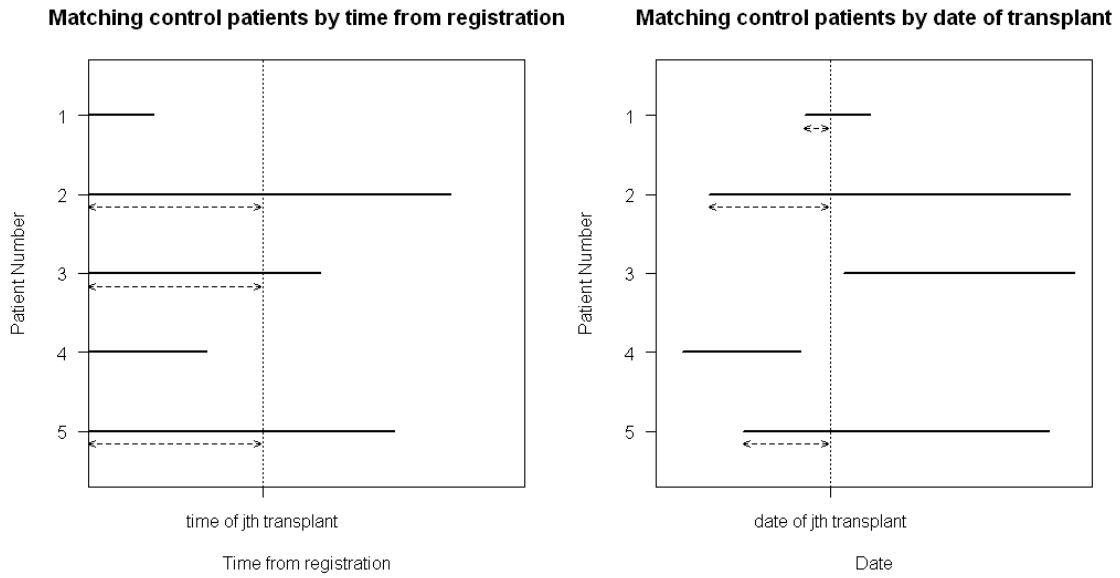

Figure 6.1: Plots showing possible differences between patients included in control groups when matching by time from registration to transplant and by date of transplant.

experiment we need to measure from a time different from time of registration. We use time from the date of the jth transplant and include previous time spent on the waiting list as a covariate in our model.

Therefore we need to define some new variables to be used when fitting a model that matches by date. These include $\text{DATE}_{T_i}, \text{DATE}_{D_i}$ and DATE_{C_i}, which are the date of transplant, date of death and date of censoring, respectively, for patient i. From these we define DATE_{Y_i}, which is the earliest date out of $\text{DATE}_{T_i}, \text{DATE}_{D_i}$ and DATE_{C_i}. We also require DATE_{A_i} which is the date that the ith patient becomes active on the waiting list. A variable is also needed that gives the amount of time patient i spent on the waiting list before the date of the jth transplant. We define this to be p_{ij}.

To show whether a patient is included in a particular experiment when matching by time from registration, an experiment entry indicator will be defined as,

$$e_{ij} = I\{\min(T_i, Y_i) \geq t_j, u_i(t_j) = u_j\}$$

for the ith patient with respect to the jth experiment. Here u_j is the UKELD score for the patient undergoing the jth transplant and $u_i(t_j)$ is the UKELD score for the ith patient at time t_j.

To show whether a patient is included in a particular experiment when matching by date, the experiment entry indicator will be defined as,

$$e_{ij} = I\{\text{DATE}_{T_j} \in [\text{DATE}_{A_i}, \text{DATE}_{Y_i}], u_i(t_j) = u_j\}$$

for the ith patient with respect to the jth experiment.

We will treat patients who are subsequently transplanted or removed from the waiting list in the same way for both methods. Patients will be censored from an experiment if they were to receive a transplant. This is because they will have triggered an experiment of their own and will no longer be contributing to mortality on the waiting list. Ideally we would follow up removals after their time of removal. However this information is not available to us, we only know whether they were removed because their condition had deteriorated or not. Therefore we will assume that any that were removed due to deteriorating condition would have died shortly after, so we assume they died on the date of removal. For individuals who were removed for other reasons, we will censor them on the date of removal, as we have have no further information about their expected survival.

Considering each individual experiment, there will be an "experimental" group and a "control" group. The patient j, who received the transplant that triggered the experiment will be the only observation in the experimental group. We can define the contributions towards the model for each type of matching method using the standard counting process format of (start, stop, event indicator).

When matching by time from registration, the patient in the experimental group will give a contribution of $(t_j, Y_j = \min(D_j, C_j), \Delta_j)$. The individuals in the corresponding control group will contribute $(t_j, \min(Y_i, T_i), \Delta_i I(D_i < T_i))$.

We can also define the contributions when matching by date, after restarting the time scale at the date of transplant. For the individual in the experimental group, this will be $(0, Y_j - p_{jj}, \Delta_j)$, where p_{jj} is the amount of time the individual in the experimental group spends on the waiting list before their transplant. Similarly, the patients in the control group will contribute $(0, \min(T_i - p_{ij}, Y_i - p_{ij}), \Delta_i I(D_i < T_i))$, where p_{ij} is the amount of time the ith individual spends on the waiting list before the jth transplant.

The model that corresponds to this proposed method of sequential stratification when finding the survival benefit for different groups of UKELD scores is

$$\lambda_{ij}^{D}(t; \boldsymbol{\theta}_0) = \lambda_{0j}^{D}(t) \exp\left\{\boldsymbol{\theta}_0^T \mathbf{Z}_{ij}\right\}, \tag{6.1}$$

where $\boldsymbol{\theta}_0 = (\boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T)^T$ and $\mathbf{Z}_{ij} = (\mathbf{Z}_{ij1}^T, \mathbf{Z}_{ij2}^T)$. Here the parameter vector of interest is $\boldsymbol{\theta}_1 = (\theta_1, \ldots, \theta_p)^T$, while $\mathbf{Z}_{ij1}$ is the $p \times 1$ covariate vector with kth component $I\{T_i = t_j\} I\{u_j \in \text{UKELD}_k\}$, where UKELD_k is the kth of the p groups of UKELD scores. The estimates of the vector $\boldsymbol{\theta}_1$ will give the UKELD-category-specific hazard ratios of post-transplant mortality versus wait-list mortality. The vector $\mathbf{Z}_{ij2}$ contains any additional adjustment covariates.

It is possible to generalise the above model to any situation. If we have a covariate X for which we wish to calculate the survival benefit of the patients at each level of the

covariate, then we could still use the model in (6.1). However, $\mathbf{Z}_{ij1}$ would become the $p \times 1$ covariate vector with kth component $I\{T_i = t_j\}I\{x_j \in X_k\}$ where X_k is the kth group of the covariate of interest. The experiment entry indicators would be redefined as

$$e_{ij} = I\{\min(T_i, Y_i) \geq t_j, x_i(t_j) = x_j\} \qquad \text{and}$$
$$e_{ij} = I\{\text{DATE}_{T_j} \in [\text{DATE}_{A_i}, \text{DATE}_{Y_i}], x_i(t_j) = x_j\},$$

when matching by time from registration and date of transplant respectively, with $x_i(t_j)$ being the value of the covariate X for the ith patient at time t_j.

6.3.1 Calculating the survival benefit of alternative transplant therapies

If we want to find the survival benefit of an alternative therapy, such as using a split liver or a liver from an extended criteria donor, relative to a standard transplant, then we need to set up the model to be used in the sequential stratification method slightly differently from that given previously. Instead of an experiment being generated by every transplant observed, only the alternative therapy transplants will initiate an experiment. This would then allow us to estimate the ratio of the hazard function of the alternative therapy relative to that of remaining on the waiting list and possibly receiving a standard transplant in the future.

We would still be able to write the model used as

$$\lambda_{ij}^D(t; \boldsymbol{\theta}_0) = \lambda_{0j}^D(t) \exp\left\{\boldsymbol{\theta}_0^T \mathbf{Z}_{ij}\right\},$$

but with $\boldsymbol{\theta}_0 = (\theta_{AT}, \boldsymbol{\theta}_0^T)$ and $\mathbf{Z}_{ij} = (Z_{ij1}, \mathbf{Z}_{ij2})$ where θ_{AT} is the parameter of interest and $Z_{ij1} = I\{T_i^{AT} = t_j\}$, where T_i^{AT} is the time that an alternative therapy transplant, such as a transplant using a split liver or an organ from an extended criteria donor, occurs for the ith patient. To be included as a control for an experiment, a patient would need to be on the waiting list at the time of transplant, so the experiment entry indicators would be

$$e_{ij} = I\{\min(T_i, Y_i) \geq t_j\} \qquad \text{and}$$
$$e_{ij} = I\{\text{DATE}_{T_j} \in [\text{DATE}_{A_i}, \text{DATE}_{Y_i}]\},$$

when matching by time from registration and date of transplant respectively. If necessary, additional constraints could be placed on the patients included as controls for a experiment so that only patients comparable to the patient who initiates the experiment are considered.

6.4 Estimating the Weight Function

The contributions of all subjects will need to be weighted to adjust for the bias introduced by the dependent censoring of transplanted patients. This will be done using weights that

are similar to inverse probability of censoring weights, which are described in detail in Section 3.1. The probability of remaining untransplanted from time t_j to time t will be given by an estimate of

$$G_{ij}(t) = \exp\left\{-\int_{t_j}^{t} d\Lambda_i^T(s)\right\}, \qquad t > t_j,$$

which can be expressed as the ratio $\frac{G_i(t)}{G_i(t_j)}$ where

$$G_i(t) = \exp\left\{-\int_{0}^{t} d\Lambda_i^T(s)\right\},$$

is the survival function for time to transplantation. The inverse of $G_{ij}(t)$ is used in the weight function and this is equivalent to using the unstabilised weights defined in Section 3.1. However the probability of remaining untransplanted starts from time t_j instead of time 0. This is why the inverse of $G_{ij}(t)$ has $G_i(t_j)$ as the numerator rather than 1.

Liver transplantation will be assumed to follow the proportional hazards model,

$$\lambda_i^T(t|\boldsymbol{\alpha}_0) = \lambda_0^T(t)\exp\left\{\boldsymbol{\alpha}_0^T\mathbf{V}_i(t)\right\}, \tag{6.2}$$

where $\lambda_0^T(t)$ is the baseline transplant hazard. This model implies that the transplant hazard for an individual only depends on the current values of vector of covariates. This is realistic as waiting list priority will be given to those who seem sickest at a particular time, as indicated by current covariate values, not their historic values.

The Cox proportional hazards model will be used to estimate the parameter vector $\boldsymbol{\alpha}_0$. The covariates that will be included in this model are all the covariates that are to be included in the model for time to death.

We will use a Kaplan-Meier estimate that has been extended to include the covariates used in the model for liver transplantation given in (6.2) as an estimate for $G_{ij}(t)$. It is not clear whether Schaubel et al (2009b) use this estimate. It is possible that they may have used $\hat{S}_i(t) = \left\{\hat{S}_0(t)\right\}^{\exp(\hat{\boldsymbol{\alpha}}'\mathbf{V}_i(t))}$ to find the estimates of the survival functions used in $G_{ij}(t)$. However, this should not be used when we have time-dependent covariates in the model.

Although we do not know which estimate for $G_{ij}(t)$ was used by Schaubel et al. (2009b), in a different paper (Zhang and Schaubel (2010)) that used similar weights, the estimate $\hat{S}_i(t) = \left\{\hat{S}_0(t)\right\}^{\exp(\hat{\boldsymbol{\alpha}}'\mathbf{V}_i(t))}$ was used with time-dependent covariates. So, it is likely that this may have been used in Schaubel et al. (2009b) as well.

Now using this model we can calculate estimates of the weighted risk set indicators,

$$\hat{W}_{ij}(t;\boldsymbol{\alpha}_0) = R_{ij}(t)\hat{G}_{ij}(t)^{-I(T_i>t_j)},$$

where

$$R_{ij}(s) = R_i(s)\{I(T_i = t_j) + I(T_i > t_j)I(T_i > s)\}$$

which is the risk set indicator for the jth experiment and $\hat{\boldsymbol{\alpha}}_0$ is the vector of parameter estimates from the Cox model for transplantation. This modifies the ordinary risk set at time s to include only those who were at risk when the experiment was initiated and have not since crossed over into the transplanted set.

6.5 Deriving an Estimating Equation

It is now necessary to derive an equation that can be used to obtain estimates for the parameters of interest. The approach used here will be slightly different from that outlined in Schaubel et al(2009b), but it is equally valid when using a counting process format for the data.

The profile likelihood for our transplantation data will be given by

$$\prod_{j=1}^{N_T}\prod_{i=1}^{n}\prod_{t\geq t_j}\left\{\frac{e_{ij}W_{ij}(t;\boldsymbol{\alpha}_0)\exp\{\boldsymbol{\theta}^T\mathbf{Z}_{ij}\}}{\sum_i e_{ij}W_{ij}(t;\boldsymbol{\alpha}_0)\exp\{\boldsymbol{\theta}^T\mathbf{Z}_{ij}\}}\right\}^{e_{ij}W_{ij}(t;\boldsymbol{\alpha}_0)dN_i(t)}.$$

From this we can easily obtain the log profile likelihood, which is given by

$$\ell(\boldsymbol{\theta},\boldsymbol{\alpha}_0) = \sum_{j=1}^{N_T}\sum_{i=1}^{n}\int_{t_j}^{\tau} e_{ij}W_{ij}(t;\boldsymbol{\alpha}_0)\Bigg[\log(e_{ij}W_{ij}(t)) + \boldsymbol{\theta}^T\mathbf{Z}_{ij} - \log\left(\sum_i e_{ij}W_{ij}(t)\exp\{\boldsymbol{\theta}^T\mathbf{Z}_{ij}\}\right)\Bigg]dN_i(t)$$

where $\tau = \max\{X_1,\ldots,X_n\}$. The estimating equation can be obtained using the score function which can be found by differentiating the log profile likelihood with respect to θ. This gives

$$\mathbf{U}(\boldsymbol{\theta},\boldsymbol{\alpha}_0) = \sum_{j=1}^{N_T}\sum_{i=1}^{n}\int_{t_j}^{\tau} e_{ij}W_{ij}(s;\boldsymbol{\alpha}_0)\left[\mathbf{Z}_{ij} - \frac{\sum_i e_{ij}W_{ij}(s;\boldsymbol{\alpha}_0)\mathbf{Z}_{ij}\exp(\boldsymbol{\theta}^T\mathbf{Z}_{ij})}{\sum_i e_{ij}W_{ij}(s;\boldsymbol{\alpha}_0)\exp(\boldsymbol{\theta}^T\mathbf{Z}_{ij})}\right]dN_i(s),$$

which can be written more compactly if we let

$$S_j^{(d)}(s;\boldsymbol{\theta},\boldsymbol{\alpha}_0) = n^{-1}\sum_{i=1}^{n} e_{ij}W_{ij}(s;\boldsymbol{\alpha}_0)\mathbf{Z}_{ij}^{\otimes d}\exp\{\boldsymbol{\theta}^T\mathbf{Z}_{ij}\} \quad \text{for} \quad d = 0,1,2$$

where $\mathbf{z}^{\otimes 0} = 1$ and $\mathbf{z}^{\otimes 1} = \mathbf{z}$ for any vector, $\mathbf{z}$ and

$$\mathbf{E}_j(s;\boldsymbol{\theta},\boldsymbol{\alpha}_0) = S_j^{(1)}(s;\boldsymbol{\theta},\boldsymbol{\alpha}_0)/S_j^{(0)}(s;\boldsymbol{\theta},\boldsymbol{\alpha}_0).$$

In the final estimating equation, an estimate of the weight function is needed, which was discussed in Section 6.4. Thus the final form of the necessary estimating equation is

$$\mathbf{U}(\boldsymbol{\theta}, \hat{\boldsymbol{\alpha}}_0) = \sum_{j=1}^{N_T} \sum_{i=1}^{n} \int_{t_j}^{\tau} e_{ij} \hat{W}_{ij}(s; \hat{\boldsymbol{\alpha}}_0) \{\mathbf{Z}_{ij} - \mathbf{E}_j(s; \boldsymbol{\theta}, \hat{\boldsymbol{\alpha}}_0)\} dN_i^D(s). \tag{6.3}$$

This is the form of the score equation for a weighted, stratified proportional hazards model, that can be fitted using standard statistical software packages, such as PROC PHREG in SAS. The only thing to note is that, as we are using weighted data, we should use the robust sandwich estimate of the covariance matrix to find the standard errors of the parameter estimates. This will give standard errors that tend to be slightly more conservative than those found using the inverse of the information matrix.

6.6 Results

The sequential stratification method is now applied to the Liver Registration data set so that estimates of θ_1 to θ_5 can be found. These are the parameters for the covariate vector $\mathbf{Z}_{ij1}$, which was defined in Section 6.3, which contains the indicators of whether the observations are equal to the jth transplant time and which UKELD group they belong to. However the values that are interest are $\exp(\hat{\theta}_1)$ to $\exp(\hat{\theta}_5)$. They are the ratios of the estimated hazard functions for patients who create experiments (those that receive transplants) to control patients (those that remain on the waiting list). If they have a value of less than 1 then those who receive transplants have a lower hazard of death. We can then use these values to determine the groups of UKELD scores in which the patients have the greatest survival benefit.

We did not include UKELD score as a covariate in the models because it had been used to match similar patients in the sequential stratification method. We split UKELD score into 5 groups when doing this. The boundaries for these groups were chosen by examining the 20%, 40%, 60% and 80% quantiles of the distribution of UKELD score and using similar values to these so that the groups contain roughly the same number of patients.

We included other covariates in these models: age at registration, primary liver disease category, ethnicity, serum sodium at time of registration and INR at time of registration. We also included previous time spent on the waiting list in the model where we matched patients by date. We do not present the parameter estimates for these covariates.

Tables 6.1 and 6.2 contain the hazard ratios for post-transplant mortality in contrast to mortality on the waiting list for 5 different levels of UKELD score when matching by time from registration. Also given are the 95% confidence intervals and p-values. The same results when matching by date can be seen in Tables 6.3 and 6.4.

k	UKELD Scores	$\exp\{\hat{\theta}_k\}$	95% Confidence Interval	P-value
1	$u < 50.5$	0.265	(0.177,0.398)	< 0.0001
2	$50.5 \leq u < 53.5$	0.150	(0.103,0.218)	< 0.0001
3	$53.5 \leq u < 56.5$	0.211	(0.146,0.305)	< 0.0001
4	$56.5 \leq u < 60$	0.121	(0.085,0.171)	< 0.0001
5	$u \geq 60$	0.169	(0.129,0.222)	< 0.0001

Table 6.1: UKELD category specific hazard ratios (post-transplant versus wait-list) when controls are matched using time from registration with UKELD model 1

k	UKELD Scores	$\exp\{\hat{\theta}_k\}$	95% Confidence Interval	P-value
1	$u < 50.5$	0.276	(0.195,0.393)	< 0.0001
2	$50.5 \leq u < 53.5$	0.171	(0.120,0.244)	< 0.0001
3	$53.5 \leq u < 56.5$	0.194	(0.138,0.273)	< 0.0001
4	$56.5 \leq u < 60$	0.146	(0.104,0.205)	< 0.0001
5	$u \geq 60$	0.150	(0.108,0.209)	< 0.0001

Table 6.2: UKELD category specific hazard ratios (post-transplant versus wait-list) when controls are matched using time from registration with UKELD model 2

k	UKELD Scores	$\exp\{\hat{\theta}_k\}$	95% Confidence Interval	P-value
1	$u < 50.5$	0.346	(0.230,0.519)	< 0.0001
2	$50.5 \leq u < 53.5$	0.171	(0.115,0.256)	< 0.0001
3	$53.5 \leq u < 56.5$	0.228	(0.157,0.331)	< 0.0001
4	$56.5 \leq u < 60$	0.127	(0.088,0.182)	< 0.0001
5	$u \geq 60$	0.206	(0.153,0.277)	< 0.0001

Table 6.3: UKELD category specific hazard ratios (post-transplant versus wait-list) when controls are matched using date of transplant with UKELD model 1

k	UKELD Scores	$\exp\{\hat{\theta}_k\}$	95% Confidence Interval	P-value
1	$u < 50.5$	0.322	(0.221,0.468)	< 0.0001
2	$50.5 \leq u < 53.5$	0.193	(0.134,0.277)	< 0.0001
3	$53.5 \leq u < 56.5$	0.184	(0.125,0.270)	< 0.0001
4	$56.5 \leq u < 60$	0.156	(0.108,0.224)	< 0.0001
5	$u \geq 60$	0.134	(0.091,0.196)	< 0.0001

Table 6.4: UKELD category specific hazard ratios (post-transplant versus wait-list) when controls are matched using date of transplant with UKELD model 2

Tables 6.1 to 6.4 show that generally UKELD groups 4 and 5 have the lowest hazard ratios, although occasionally group 2 has a fairly low hazard ratio. This means they have the greatest difference between waiting list and post-transplant mortality, with post-transplant survival expected to be much greater than waiting list survival. In terms of survival benefit, this means that the patients in UKELD groups 4 and 5 generally have the highest survival benefit from liver transplantation, under these models. Also, UKELD group 1 always has the highest hazard ratio, which means the patients with the lowest UKELD scores have the lowest survival benefit.

These results make intuitive sense as UKELD score is a good predictor of mortality, we know that as the UKELD score increases, the expected survival on the waiting list decreases. So the contrast between waiting list and post-transplant mortality should increase as long as post-transplant mortality does not also decrease rapidly as UKELD score increases.

The results also suggest that the transplants being carried out on patients with high UKELD scores are not futile, as they can expect a significant improvement in their survival after transplant. However, we should be aware that the data we are considering are observational data. So, any patients with high UKELD scores that receive liver transplants have been deemed as suitable for transplant by a surgeon. Therefore the result we see here may be a consequence of this selection bias in our data.

In addition to these fitted models, we also produced bootstrap confidence intervals to assess the robustness of the results from these models. We carried out B bootstrap replications, each time sampling from our dataset with replacement and then applying the sequential stratification method to the new dataset. The distributions of the bootstrap estimates for $\exp(\theta_1)$ to $\exp(\theta_5)$ were then examined and the 2.5 and 97.5 percentiles of the distributions were used to find 95% confidence intervals for these parameter estimates.

Tables 6.5 and 6.6 give the bootstrap confidence intervals when matching by time from registration using UKELD model 1 and UKELD model 2 respectively. The bootstrap confidence intervals for the same models, but matching by date instead, can be seen in Tables 6.7 and 6.8. The histograms of the bootstrap estimates for $\exp(\theta_1)$ to $\exp(\theta_5)$ for each of these models were examined to ensure that the estimates were approximately normally distributed. The histograms were roughly bell shaped, with the majority being roughly symmetric, although there was occasionally some skewness, particularly in the distributions of the bootstrap estimates for UKELD group 1.

There is considerable overlap between the bootstrap confidence intervals in each of these tables. However, the intervals for $\exp(\hat{\theta}_1)$ do tend to be much wider than those for the other four parameters. These results suggest that a UKELD score of greater than 50.5 is all that is needed for patients to receive a significant benefit from transplantation.

Hazard Ratio	Bootstrap 95% Confidence Interval
$\exp(\hat{\theta}_1)$	(0.126,0.946)
$\exp(\hat{\theta}_2)$	(0.057,0.370)
$\exp(\hat{\theta}_3)$	(0.088,0.404)
$\exp(\hat{\theta}_4)$	(0.056,0.208)
$\exp(\hat{\theta}_5)$	(0.085,0.301)

Table 6.5: Table showing 95% confidence intervals for the UKELD group specific hazard ratios based on percentiles of the distribution of bootstrap estimates when B=1,000 and matching by time from registration. The UKELD model being used here is UKELD model 1.

Hazard Ratio	Bootstrap 95% Confidence Interval
$\exp(\hat{\theta}_1)$	(0.137,0.959)
$\exp(\hat{\theta}_2)$	(0.070,0.413)
$\exp(\hat{\theta}_3)$	(0.087,0.371)
$\exp(\hat{\theta}_4)$	(0.070,0.246)
$\exp(\hat{\theta}_5)$	(0.073,0.271)

Table 6.6: Table showing 95% confidence intervals for the UKELD group specific hazard ratios based on percentiles of the distribution of bootstrap estimates when B=1,000 and matching by time from registration. The UKELD model being used here is UKELD model 2.

Hazard Ratio	Bootstrap 95% Confidence Interval
$\exp(\hat{\theta}_1)$	(0.151,0.976)
$\exp(\hat{\theta}_2)$	(0.086,0.354)
$\exp(\hat{\theta}_3)$	(0.118,0.423)
$\exp(\hat{\theta}_4)$	(0.055,0.214)
$\exp(\hat{\theta}_5)$	(0.096,0.319)

Table 6.7: Table showing 95% confidence intervals for the UKELD group specific hazard ratios based on percentiles of the distribution of bootstrap estimates when B=1,000 and matching by date. The UKELD model being used here is UKELD model 1.

Hazard Ratio	Bootstrap 95% Confidence Interval
$\exp(\hat{\theta}_1)$	(0.159,0.975)
$\exp(\hat{\theta}_2)$	(0.097,0.390)
$\exp(\hat{\theta}_3)$	(0.100,0.331)
$\exp(\hat{\theta}_4)$	(0.076,0.278)
$\exp(\hat{\theta}_5)$	(0.064,0.220)

Table 6.8: Table showing 95% confidence intervals for the UKELD group specific hazard ratios based on percentiles of the distribution of bootstrap estimates when B=1,000 and matching by date. The UKELD model being used here is UKELD model 2.

However, different numbers of intervals and boundaries for these intervals need to be investigated before we can be certain of this.

When the confidence intervals in Tables 6.5 to 6.8 are compared to those in Tables 6.1 to 6.4, we see that the bootstrap confidence intervals are always wider than the those from the fitted models. However, for the final two UKELD groups, the bootstrap confidence intervals tend to be fairly close to the confidence intervals from the fitted models. This suggests that the results for these two groups are fairly robust to changes in the individuals included in the data set.

Figure 6.2 gives a graphical representation of all the results for the models considered so far. For each UKELD group we show the hazard ratios and confidence intervals from the fitted models using each of the UKELD models, alongside the bootstrap confidence intervals for the same models. This allows easy comparison of the results from each of the models much more easily. The horizontal line on the plots corresponds to a hazard ratio of value 1. If a confidence interval crosses this line then the difference between waiting list and post-transplant mortality is not considered to be significant.

Generally, we see that the hazard ratios and the upper limits of the confidence intervals tend to decrease as the UKELD score increases. However, under UKELD model 1 there is a slight increase in the hazard ratios and 95% confidence intervals for UKELD group 5 compared to UKELD group 4. This increase in the 95% confidence interval means there is more uncertainty about the estimate of the hazard ratio for UKELD group 5.

From the plots, we see that the bootstrap confidence intervals for UKELD group 1 are much wider than those for any of the other groups. They are also close to including a hazard ratio of 1, suggesting that difference between waiting list and post-transplant mortality is only just significant.

It is also much easier to see just how much overlap there is between the confidence intervals, particularly between UKELD groups 2 and 3 and UKELD groups 4 and 5.

For this particular data set, the results suggest that there is very little difference be-

tween using time from registration to match controls and using date of transplant to match the control patients. However, if the survival benefit of a fairly new alternative transplant therapy is being calculated, where we could expect to see a noticeable improvement in the survival of patients over time, then matching control patients by date of transplant may give more realistic results.

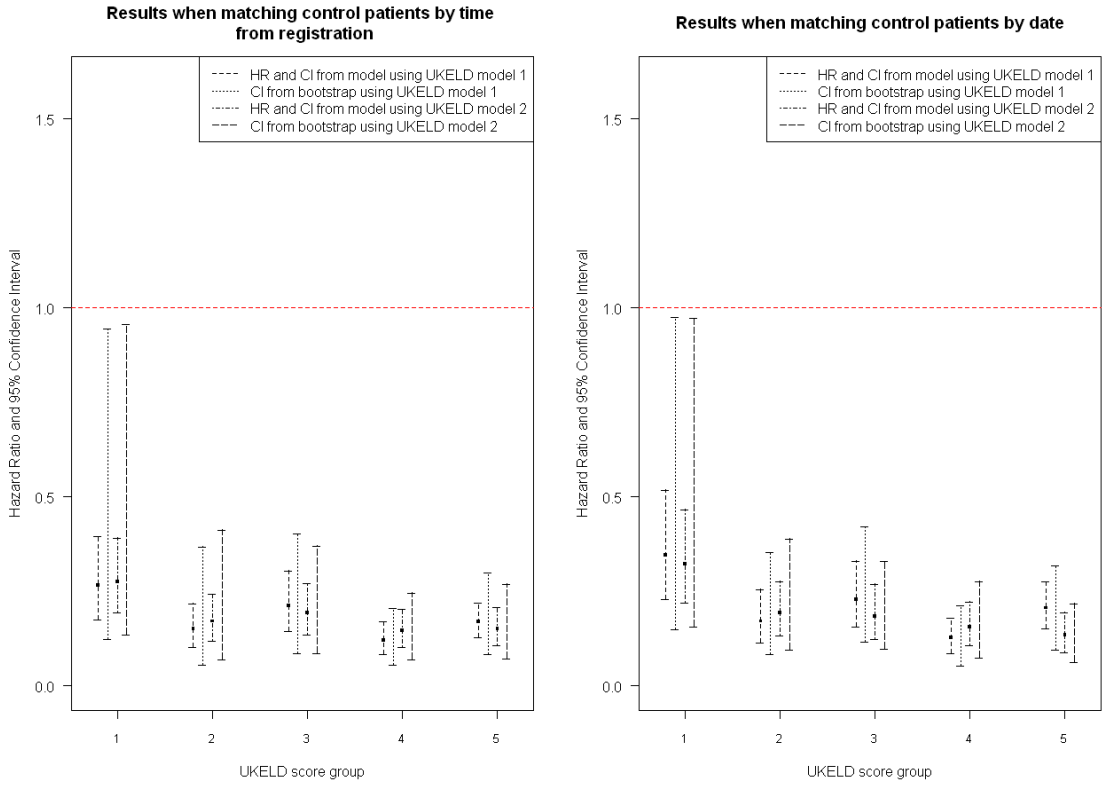

Figure 6.2: Plots showing the hazard ratios and 95% confidence intervals of post-transplant mortality versus waiting list mortality from both the model and bootstrap when matching control patients by time from registration and when matching control patients by date

6.7 Using Additional Criteria when Matching Control Patients

So far we have included as controls for the experiments any patients who were still on the waiting list either at the time of transplantation or on the date of the transplant. However, in reality, only patients who are deemed suitable for the donor organ would be considered for transplant. Thus we should incorporate some of these additional criteria into our model, so that we only use patients that are comparable to the experiment generating patient as controls.

One of the most important criteria when deciding if a patient is suitable for a transplant is whether he is blood group compatible with the donor of the organ. If the donor has blood group O, then any patient can have the organ. If the donor has blood group A, then the organ can only be given to a patient with blood group A or AB. If the donor has blood group B, then the organ can only be given to a patient with blood group B or AB. If the donor has blood group AB, then only patients who also have blood group AB can receive the organ.

If we only include patients who are blood group compatible when matching by time from registration, then the experiment entry indicator would be

$$\begin{aligned} e_{ij} = I\{ \min(T_i, Y_i) \geq t_j, u_i(t_j) = u_j, rbg_i = A \text{ or } AB \text{ if } dbg_j = A, \\ rbg_i = B \text{ or } AB \text{ if } dbg_j = B, rbg_i = AB \text{ if } dbg_j = AB\} \end{aligned}$$

for the ith patient with respect to the jth experiment. Here rbg_i is the blood group of the ith potential recipient and dbg_j is the blood group of the donor of the organ that is used in the jth transplant. Similarly the experiment entry indicator when matching controls by date of transplant and blood group would become

$$\begin{aligned} e_{ij} = I\{\text{DATE}_{T_j} \in [\text{DATE}_{A_i}, \text{DATE}_{Y_i}], u_i(t_j) = u_j, rbg_i = A \text{ or } AB \\ \text{if } dbg_j = A, rbg_i = B \text{ or } AB \text{ if } dbg_j = B, rbg_i = AB \text{ if } dbg_j = AB\} \end{aligned}$$

for the ith patient with respect to the jth experiment.

As blood group compatibility is so important when choosing a recipient for a donor organ, it should be included in the final model that we use to find the survival benefit for the different groups of UKELD scores.

Another criterion that is considered when selecting recipients for a donor organ is the difference between the weight of the recipient and the weight of the donor. If this difference is too large, then the donor organ could be the wrong size. Therefore patients are usually only considered if their weight is within 10kg of the weight of the donor. However, this is not strictly adhered to, as we found by looking at the differences in

weights for the patients who were transplanted. So it is not as important to include this in the model as the previous criterion. For this reason we will incorporate it in a model that already considers blood group compatible patients, and compare the results to a model that incorporates only blood group compatibility.

If we include patients who are both blood group and weight compatible when matching by time from registration, then the experiment entry indicator would be

$$e_{ij} = I\{\min(T_i, Y_i) \geq t_j, u_i(t_j) = u_j, rbg_i = A \text{ or } AB \text{ if } dbg_j = A, rbg_i = B \\ \text{or } AB \text{ if } dbg_j = B, rbg_i = AB \text{ if } dbg_j = AB, |dw_j = rw_i| \leq 10kg\}$$

for the ith patient with respect to the jth experiment. Here rw_i is the weight of the ith patient and dw_j is the weight of the donor of the organ that is used in the jth transplant. Similarly the experiment entry indicator when matching controls by date of transplant, blood group and weight would become

$$e_{ij} = I\{\text{DATE}_{T_j} \in [\text{DATE}_{A_i}, \text{DATE}_{Y_i}], u_i(t_j) = u_j, rbg_i = A \text{ or } AB \text{ if } dbg_j = A, \\ rbg_i = B \text{ or } AB \text{ if } dbg_j = B, rbg_i = AB \text{ if } dbg_j = AB, |dw_j = rw_i| \leq 10kg\}$$

for the ith patient with respect to the jth experiment.

6.7.1 Results

Here we apply the sequential stratification method to our data set but also incorporate some of the additional criteria described in the previous section. Firstly, we consider models that match by time from registration and blood group for both UKELD models. The results for these models are given in Tables 6.9 and 6.10. Then models that match by time from registration, blood group and weight are presented, again using each of the UKELD models considered. Tables 6.11 and 6.12 contain the results for these models.

The decreasing patterns in the hazard ratios and the limits of the confidence intervals are even more evident in these tables than in previous results. Again those in UKELD groups 4 and 5 have the largest survival benefit and those in group 1 tend to have the lowest survival benefit.

We have also produced 95% confidence intervals based on the percentiles of the distribution of the bootstrap estimates for $\exp(\theta_1)$ to $\exp(\theta_5)$. The aim here is to produce some more robust confidence intervals for the hazard ratios that are of interest. These are given in Tables 6.13 to 6.16.

As seen previously, the bootstrap confidence intervals are always wider than those from the fitted model. There is also much overlap between the bootstrap confidence intervals given in each table. However, the confidence intervals do tend to get tighter for each successive UKELD group and the upper limit of the confidence interval tends to decrease.

k	UKELD Scores	$\exp\{\hat{\theta}_k\}$	95% Confidence Interval	P-value
1	$u < 50.5$	0.264	(0.173,0.402)	< 0.0001
2	$50.5 \leq u < 53.5$	0.234	(0.154,0.354)	< 0.0001
3	$53.5 \leq u < 56.5$	0.194	(0.132,0.287)	< 0.0001
4	$56.5 \leq u < 60$	0.132	(0.093,0.189)	< 0.0001
5	$u \geq 60$	0.113	(0.085,0.152)	< 0.0001

Table 6.9: UKELD category specific hazard ratios (post-transplant versus wait-list) using UKELD model 1, when controls are matched using time from registration and blood group.

k	UKELD Scores	$\exp\{\hat{\theta}_k\}$	95% Confidence Interval	P-value
1	$u < 50.5$	0.259	(0.179,0.376)	< 0.0001
2	$50.5 \leq u < 53.5$	0.278	(0.189,0.407)	< 0.0001
3	$53.5 \leq u < 56.5$	0.191	(0.135,0.270)	< 0.0001
4	$56.5 \leq u < 60$	0.162	(0.115,0.227)	< 0.0001
5	$u \geq 60$	0.104	(0.072,0.149)	< 0.0001

Table 6.10: UKELD category specific hazard ratios (post-transplant versus wait-list) using UKELD model 2, when controls are matched using time from registration and blood group.

k	UKELD Scores	$\exp\{\hat{\theta}_k\}$	95% Confidence Interval	P-value
1	$u < 50.5$	0.358	(0.234,0.548)	< 0.0001
2	$50.5 \leq u < 53.5$	0.276	(0.178,0.428)	< 0.0001
3	$53.5 \leq u < 56.5$	0.178	(0.117,0.271)	< 0.0001
4	$56.5 \leq u < 60$	0.139	(0.097,0.200)	< 0.0001
5	$u \geq 60$	0.109	(0.079,0.150)	< 0.0001

Table 6.11: UKELD category specific hazard ratios (post-transplant versus wait-list) using UKELD model 1, when controls are matched using time from registration, blood group and weight.

k	UKELD Scores	$\exp\{\hat{\theta}_k\}$	95% Confidence Interval	P-value
1	$u < 50.5$	0.353	(0.240,0.518)	< 0.0001
2	$50.5 \leq u < 53.5$	0.349	(0.233,0.523)	< 0.0001
3	$53.5 \leq u < 56.5$	0.182	(0.126,0.263)	< 0.0001
4	$56.5 \leq u < 60$	0.160	(0.113,0.227)	< 0.0001
5	$u \geq 60$	0.107	(0.073,0.157)	< 0.0001

Table 6.12: UKELD category specific hazard ratios (post-transplant versus wait-list) using UKELD model 2, when controls are matched using time from registration, blood group and weight.

There are two bootstrap confidence intervals for UKELD group 1 that suggest the estimated hazard ratio is not as significant as was suggested under the fitted model. In Table 6.15, the confidence interval for UKELD group 1 includes the value 1, which suggests there is no significant difference between waiting list and post-transplant mortality here. In Table 6.16, the confidence interval for UKELD group 1 does not include the value 1, but the upper limit of the interval is close to it, which suggests the estimated hazard ratio here is only just significant.

Hazard Ratio	Bootstrap 95% Confidence Interval
$\exp(\hat{\theta}_1)$	(0.119,0.816)
$\exp(\hat{\theta}_2)$	(0.087,0.542)
$\exp(\hat{\theta}_3)$	(0.085,0.373)
$\exp(\hat{\theta}_4)$	(0.061,0.251)
$\exp(\hat{\theta}_5)$	(0.056,0.253)

Table 6.13: Table showing 95% confidence intervals for the UKELD group specific hazard ratios based on percentiles of the distribution of bootstrap estimates when B=1,000 and matching by time from registration and blood group. The UKELD model being used here is UKELD model 1.

Hazard Ratio	Bootstrap 95% Confidence Interval
$\exp(\hat{\theta}_1)$	(0.121,0.794)
$\exp(\hat{\theta}_2)$	(0.116,0.600)
$\exp(\hat{\theta}_3)$	(0.081,0.358)
$\exp(\hat{\theta}_4)$	(0.079,0.300)
$\exp(\hat{\theta}_5)$	(0.046,0.238)

Table 6.14: Table showing 95% confidence intervals for the UKELD group specific hazard ratios based on percentiles of the distribution of bootstrap estimates when B=1,000 and matching by time from registration and blood group. The UKELD model being used here is UKELD model 2.

Figure 6.3 gives a graphical representation of all the results for models that match controls by time from registration and use additional criteria when matching. For each UKELD group we show the hazard ratios and confidence intervals from the fitted models using each of the UKELD models, alongside the bootstrap confidence intervals for the same models. As before, the horizontal line on the plots corresponds to a hazard ratio of value 1.

These plots provide a summary of all the results in Tables 6.9 to 6.16, they can be used to interpret the results of models that match controls by time from registration and use

Hazard Ratio	Bootstrap 95% Confidence Interval
$\exp(\hat{\theta}_1)$	(0.154,1.046)
$\exp(\hat{\theta}_2)$	(0.107,0.623)
$\exp(\hat{\theta}_3)$	(0.065,0.359)
$\exp(\hat{\theta}_4)$	(0.064,0.258)
$\exp(\hat{\theta}_5)$	(0.053,0.261)

Table 6.15: Table showing 95% confidence intervals for the UKELD group specific hazard ratios based on percentiles of the distribution of bootstrap estimates when B=1,000 and matching by time from registration, blood group and weight. The UKELD model being used here is UKELD model 1.

Hazard Ratio	Bootstrap 95% Confidence Interval
$\exp(\hat{\theta}_1)$	(0.158,0.949)
$\exp(\hat{\theta}_2)$	(0.142,0.756)
$\exp(\hat{\theta}_3)$	(0.072,0.347)
$\exp(\hat{\theta}_4)$	(0.073,0.291)
$\exp(\hat{\theta}_5)$	(0.046,0.256)

Table 6.16: Table showing 95% confidence intervals for the UKELD group specific hazard ratios based on percentiles of the distribution of bootstrap estimates when B=1,000 and matching by time from registration, blood group and weight. The UKELD model being used here is UKELD model 2.

at least one of the additional criteria when matching. We see that the downward trend in the hazard ratios and the upper limits of the confidence intervals is more pronounced here than in Figure 6.2. We can also see that there is a lot of overlap between the confidence intervals, particularly between groups 1 and 2 and groups 3 and 4.

In the plot for models that match controls using time from registration and blood group, we see that the bootstrap confidence intervals for UKELD group 1 are not as wide as those in Figure 6.2. This would be the model that is recommended for use as it incorporates what is considered to be the most important additional criteria without the size of the control groups appearing to be too greatly reduced. It does not matter which of the UKELD models is used as the results seem to be fairly robust to the choice of UKELD model.

From the plot for models that match controls by time from registration, blood group and weight, we can see that the bootstrap confidence interval for UKELD group 1 using UKELD model 1 suggests the difference between waiting list and post-transplant mortality is not significant. It is also possible to see that for UKELD groups 1 and 2 there is much more uncertainty in the estimates. This could be due to a reduction in the size of the control groups caused by using the weight matching criterion.

We now apply the sequential stratification method with the additional criteria to models that use date of transplant to match control patients. Firstly, we considered models that matched by date and blood group for both UKELD models. The results for these models are given in Tables 6.17 and 6.18. Then models that matched by date, blood group and weight are presented, again using each of the UKELD models considered. Tables 6.19 and 6.20 contain the results for these models.

The patterns that we see in the results in Tables 6.17 to 6.20 are not as clear as those in the results for matching controls by time from registration. UKELD groups 4 and 5 still tend to have the lowest hazard ratios and therefore the greatest survival benefit, although this is not always the case. For example, in Table 6.18 the hazard ratio for UKELD group 3 is lower than that for UKELD group 4. The groups with the lower UKELD scores still have the higher hazard ratios and so the lower values of survival benefit. However, here it is often UKELD group 2 that has the highest hazard ratio rather than UKELD group 1. However, there is a substantial overlap in all of the confidence intervals.

Again we have produced bootstrap confidence intervals for these models, and they are given in Tables 6.21 to 6.24. The confidence intervals in these tables are all wider than those for the fitted model, as expected from previous results. As before, there is considerable overlap between the confidence intervals shown in each table. But the confidence intervals do get tighter and have lower upper limits for the models that use UKELD model 2. For the models that use UKELD model 1, there seems to be more uncertainty about the

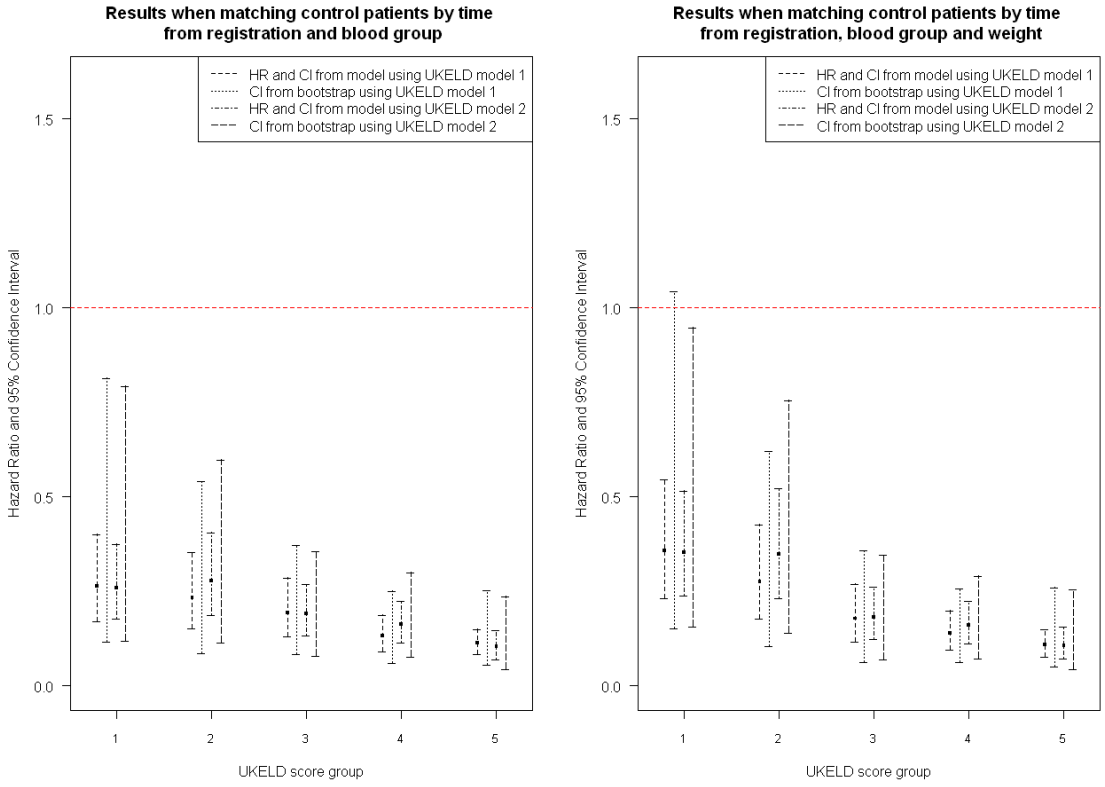

Figure 6.3: Plots showing the hazard ratios and 95% confidence intervals of post-transplant mortality versus waiting list mortality from both the model and bootstrap when matching control patients by time from registration and using additional criteria

k	UKELD Scores	$\exp\{\hat{\theta}_k\}$	95% Confidence Interval	P-value
1	$u < 50.5$	0.288	(0.190,0.437)	< 0.0001
2	$50.5 \leq u < 53.5$	0.272	(0.171,0.432)	< 0.0001
3	$53.5 \leq u < 56.5$	0.186	(0.119,0.291)	< 0.0001
4	$56.5 \leq u < 60$	0.154	(0.105,0.228)	< 0.0001
5	$u \geq 60$	0.157	(0.113,0.218)	< 0.0001

Table 6.17: UKELD category specific hazard ratios (post-transplant versus wait-list) using UKELD model 1, when controls are matched using date of transplant and blood group

k	UKELD Scores	$\exp\{\hat{\theta}_k\}$	95% Confidence Interval	P-value
1	$u < 50.5$	0.264	(0.178,0.390)	< 0.0001
2	$50.5 \leq u < 53.5$	0.274	(0.179,0.419)	< 0.0001
3	$53.5 \leq u < 56.5$	0.191	(0.129,0.283)	< 0.0001
4	$56.5 \leq u < 60$	0.208	(0.142,0.305)	< 0.0001
5	$u \geq 60$	0.113	(0.074,0.173)	< 0.0001

Table 6.18: UKELD category specific hazard ratios (post-transplant versus wait-list) using UKELD model 2, when controls are matched using date of transplant and blood group

k	UKELD Scores	$\exp\{\hat{\theta}_k\}$	95% Confidence Interval	P-value
1	$u < 50.5$	0.343	(0.213,0.554)	< 0.0001
2	$50.5 \leq u < 53.5$	0.392	(0.240,0.637)	< 0.0001
3	$53.5 \leq u < 56.5$	0.278	(0.178,0.436)	< 0.0001
4	$56.5 \leq u < 60$	0.154	(0.102,0.232)	< 0.0001
5	$u \geq 60$	0.176	(0.125,0.249)	< 0.0001

Table 6.19: UKELD category specific hazard ratios (post-transplant versus wait-list) using UKELD model 1, when controls are matched using date of transplant, blood group and weight

k	UKELD Scores	$\exp\{\hat{\theta}_k\}$	95% Confidence Interval	P-value
1	$u < 50.5$	0.322	(0.207,0.500)	< 0.0001
2	$50.5 \leq u < 53.5$	0.383	(0.248,0.591)	< 0.0001
3	$53.5 \leq u < 56.5$	0.223	(0.150,0.331)	< 0.0001
4	$56.5 \leq u < 60$	0.217	(0.146,0.321)	< 0.0001
5	$u \geq 60$	0.157	(0.104,0.237)	< 0.0001

Table 6.20: UKELD category specific hazard ratios (post-transplant versus wait-list) using UKELD model 2, when controls are matched using date of transplant, blood group and weight

hazard ratio for UKELD group 5, than that for UKELD group 4.

Most of the confidence intervals here suggest that there is not a significant difference between waiting list and post-transplant mortality for UKELD group 1. The confidence interval for UKELD group 1 in Table 6.22 suggests that it is only just significant. The confidence intervals for this group are so wide here which suggests that the control groups are too small to make precise inferences about the hazard ratio.

Hazard Ratio	Bootstrap 95% Confidence Interval
$\exp(\hat{\theta}_1)$	(0.114,1.039)
$\exp(\hat{\theta}_2)$	(0.102,0.624)
$\exp(\hat{\theta}_3)$	(0.071,0.410)
$\exp(\hat{\theta}_4)$	(0.059,0.299)
$\exp(\hat{\theta}_5)$	(0.070,0.343)

Table 6.21: Table showing 95% confidence intervals for the UKELD group specific hazard ratios based on percentiles of the distribution of bootstrap estimates when B=1,000 and matching by date and blood group. The UKELD model being used here is UKELD model 1.

Hazard Ratio	Bootstrap 95% Confidence Interval
$\exp(\hat{\theta}_1)$	(0.115,0.950)
$\exp(\hat{\theta}_2)$	(0.109,0.620)
$\exp(\hat{\theta}_3)$	(0.074,0.391)
$\exp(\hat{\theta}_4)$	(0.085,0.395)
$\exp(\hat{\theta}_5)$	(0.045,0.275)

Table 6.22: Table showing 95% confidence intervals for the UKELD group specific hazard ratios based on percentiles of the distribution of bootstrap estimates when B=1,000 and matching by date and blood group. The UKELD model being used here is UKELD model 2.

Figure 6.4 gives a graphical representation of all the results for models that match controls by date of transplant and use at least one of the additional criteria when matching. As before, we show the hazard ratios and confidence intervals from the fitted models using each of the UKELD models, alongside the bootstrap confidence intervals for the same models, for each UKELD group.

Figure 6.4 provides a summary of the results in Tables 6.17 to 6.24 and can be used to interpret the results of models that match controls by date of transplant and also use at least one of the additional criteria when matching. We see that generally there is still

Parameter	Bootstrap 95% Confidence Interval
$\exp(\hat{\theta}_1)$	(0.113,1.511)
$\exp(\hat{\theta}_2)$	(0.116,0.986)
$\exp(\hat{\theta}_3)$	(0.100,0.614)
$\exp(\hat{\theta}_4)$	(0.056,0.310)
$\exp(\hat{\theta}_5)$	(0.066,0.381)

Table 6.23: Table showing 95% confidence intervals for the UKELD group specific hazard ratios based on percentiles of the distribution of bootstrap estimates when B=1,000 and matching by date, blood group and weight. The UKELD model being used here is UKELD model 1.

Hazard Ratio	Bootstrap 95% Confidence Interval
$\exp(\hat{\theta}_1)$	(0.133,1.382)
$\exp(\hat{\theta}_2)$	(0.122,0.893)
$\exp(\hat{\theta}_3)$	(0.078,0.471)
$\exp(\hat{\theta}_4)$	(0.083,0.423)
$\exp(\hat{\theta}_5)$	(0.052,0.360)

Table 6.24: Table showing 95% confidence intervals for the UKELD group specific hazard ratios based on percentiles of the distribution of bootstrap estimates when B=1,000 and matching by date, blood group and weight. The UKELD model being used here is UKELD model 2.

a decreasing trend in the hazard ratios and the upper limits of the confidence intervals, although it is not as clear to see as in Figure 6.3.

In the plot for models that match control patients by date and blood group, the bootstrap confidence interval for UKELD group 1 using UKELD model 1 suggests that there is not a significant difference between the waiting list and post-transplant mortality. The difference is only barely significant if we consider the bootstrap confidence interval for this UKELD group using UKELD model 2. We can also see that there is much overlap between the confidence intervals for UKELD groups 1 and 2 and UKELD groups 3 and 4.

In the plot for models that match control patients by date, blood group and weight, both the bootstrap confidence intervals for UKELD group 1 suggest that there is no significant difference between waiting list and post-transplant mortality. Also the bootstrap confidence interval for UKELD group 2 under UKELD model 1 suggests that the difference here is only just significant. It is likely that the additional criteria that have been applied here have made the control groups too small, which is why we see so much uncertainty in the estimated hazard ratios.

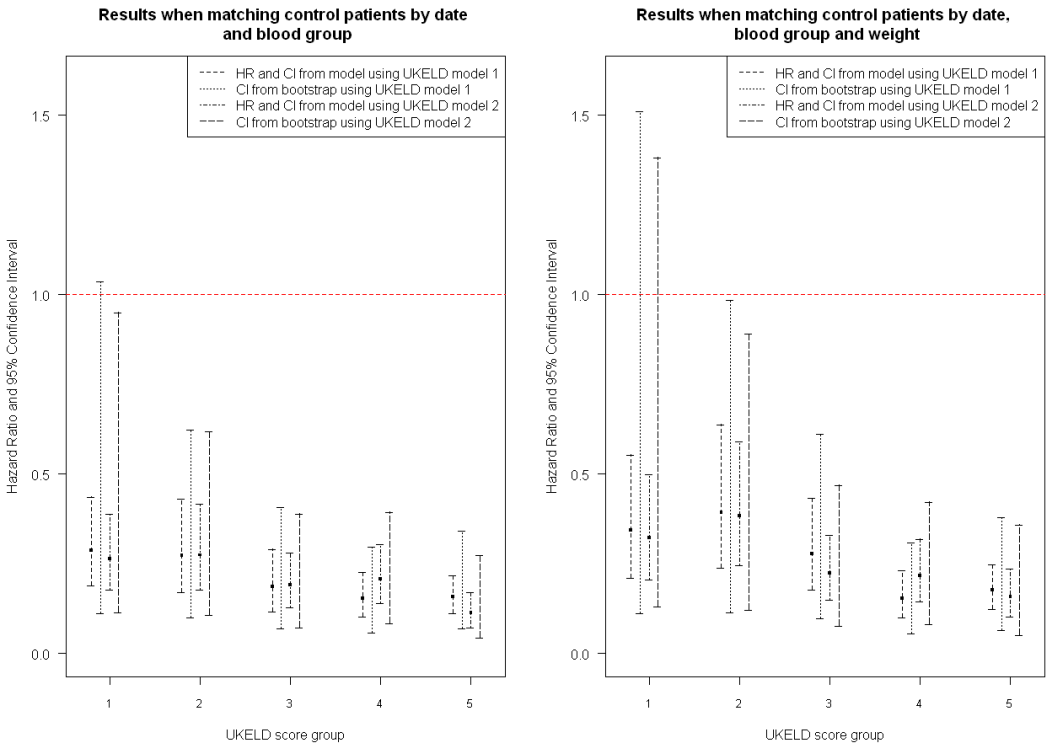

Figure 6.4: Plots showing the hazard ratios and 95% confidence intervals of post-transplant mortality versus waiting list mortality from both the model and bootstrap when matching control patients by date and using additional criteria

6.8 Summary and Recommendations

In this chapter, we describe the sequential stratification method which creates a stratum each time a patient is transplanted and compares his/her survival to those of similar candidates who were active on the waiting list at the time. This method was presented in Schaubel (2009b), although we have made a few alterations to the method.

We use this to derive the survival benefit for different UKELD score groups using covariate-adjusted hazard ratios for transplantation compared to not receiving a transplant. We found that the groups with the highest UKELD scores have the lowest hazard ratios and so have the greatest survival benefit.

We present two methods for selecting the patients that are used as comparisons for the experimental patient. The first is the one used in Schaubel (2009b), where the comparison patients are those that have been on the waiting list for at least the same amount of time as the experimental patient. We have developed the second method, where the patients used for comparison are those that are registered as active on the waiting list on the date of the transplant of the experimental patient.

Here, the results of the two methods are similar, suggesting that it does not matter which one is used. However, if there is likely to be a change in the expected survival of patients receiving a particular therapy over time, then using patients that are registered as active on the date of transplant for comparison may give more realistic results.

We also describe how the same method could be used to compare the hazard function for an alternative transplantation therapy, such as a split liver or a liver from an extended criteria donor, with the hazard function of remaining on the waiting list and possibly receiving a standard transplant at a later date.

We also considered using additional criteria when choosing patients to be included in the comparative group, which were blood group compatibility and a suitable weight relative to the weight of the donor. The results of these models showed the same trends as the results of the models without these additional criteria, so that those with the highest UKELD scores have the greatest survival benefit from liver transplantation.

However, we recommend using a model that ensures patients that are included in the comparative group are also blood group compatible, as this makes our model more realistic and there is also less uncertainty about the estimates produced by this model.

When applying the method outlined in this chapter, we must be aware that we are using observational data and therefore there may be bias in our results because of this. One particular example is selection bias. All of the patients in the data set have been chosen by clinicians for transplantation, which means they were considered suitable for the particular organ and well enough to undergo the procedure. Therefore the results from

the sequential stratification method may not be representative of all patients in the group, particularly for those with high UKELD scores as they are the most sick on the list.

Chapter 7

Discussion and Future Work

The aim of this thesis is to develop suitable methodologies for analysing data from patients on the waiting list for a liver transplant, where patients who are censored due to transplantation are suspected to be informatively censored. These methodologies should allow the survival function to be estimated as well as any significant covariates to be identified. Ultimately, they should be able to be developed into methods that can calculate other values that are of interest to NHSBT, such as survival benefit.

A detailed discussion of how this thesis meets these aims is given in Section 7.1. A summary of the main strengths and weaknesses is given in Section 7.2. Many of the methods discussed in Chapters 2 and 3 and the method developed in Chapter 4 can be applied to other situations instead of the liver transplantation setting considered in this thesis. Therefore, Section 7.3 gives recommendations on how to analyse general data with potentially informative censoring. We also explain in Section 7.4 how the methods developed in this thesis are of use to NHSBT, which provided the funding for this project. Finally, extensions of the methods developed and possible future work are discussed in Section 7.5.

7.1 Discussion

Estimators that can be used to give bounds on the estimated survival function are reviewed in Chapter 2. All these estimators are applied to the Liver Registration data set, but it is found that they give bounds that are too wide to be of use. Slud and Rubinstein (1983) and Klein and Moeschberger (1988) suggest restricting the values that the dependence parameters can take to provide tighter bounds on the estimated survival function. However, even these bounds are too wide to be useful.

Even though the bounds on the estimated survival function are not useful, these estimators can still be used to estimate the survival function if a suitable value of the dependence

parameter is specified. As we cannot identify the amount of dependence between T and C from the observed data, then this approach is not recommended. These estimators also do not allow all covariates to be incorporated, which is another reason why we would not recommend their use in practice. However, we will still discuss the properties of these estimators so that we can identify which is most suitable for use in the liver transplantation setting.

We would not recommend use of either the estimator in Fisher and Kanarek (1974) or the estimator in Slud and Rubinstein (1983), as it is not easy to specify an amount of dependence between T and C that can be interpreted easily using standard measures of dependence. All the other estimators considered in Chapter 2 use Kendall's τ to specify the amount of dependence between T and C. Use of the Fisher-Kanarek estimator is also not recommended due to some strange behaviour that can be observed when the last observation is censored.

We recommend that the copula-graphic estimator is used instead of the self-consistent estimator, when using an estimator with an assumed copula, as it is less computationally intensive. Also, it is found in a simulation study in Zheng and Klein (1994) that the self-consistent estimator has a significantly larger bias than the copula-graphic estimator.

It is not known how the other estimators in Chapter 2 compare to the copula-graphic estimator. However, as these methods cannot easily be used in practice due to the wide bounds found and the difficulties with incorporating covariates, it would not be particularly useful to identify the preferred estimator of those in Chapter 2.

The literature review is continued in Chapter 3, where methods that can incorporate covariates and are generally of more use practically are considered. These methods can be split into two categories: estimators that use models of the censoring process and sensitivity analyses.

The most widely used approaches in the literature are estimators that use a regression model for time to censoring. These estimators are known as inverse probability of censoring weighted (IPCW) estimators. These estimators are weighted versions of the standard methods, with the weights being the inverse of the probability of the individual remaining uncensored under the chosen regression model for time to censoring. This allows us to find the KM estimate of the survival function or the parameter estimates for the Cox model in the absence of any censoring. The models for time to censoring that are considered are Cox's proportional hazards model, Weibull proportional hazards model and Aalen's additive hazard model. We feel that Cox's proportional hazards model is the best model to use as it can easily incorporate time-dependent covariates and can also be fitted easily using standard software.

For IPCW estimators to be unbiased, the assumption of sequential ignorability of

censoring needs to hold. Consequently, if all the prognostic factors for both T and C are adjusted for in the model for censoring, then C is independent of T. This assumption is more restrictive than some of the other assumptions used in the methods discussed in Chapter 3. The assumption that is used in Siannis (2004), Siannis et al. (2005) and Siannis (2011) can be related to the semi-parametric model in Scharfstein and Robins (2002), which has the assumption of sequential ignorability of censoring as a special case.

However, the assumption of sequential ignorability of censoring is an intuitive choice as it seems likely that dependence between T and C would be due to shared prognostic factors. But it is possible that some of these prognostic factors are unmeasured and there would be residual dependence between T and C that is not explained by the shared factors included in the model for time to censoring. If there is residual dependence then the IPCW estimates would be biased. Although if the most significant shared prognostic factors are included in the model for time to censoring, then this bias should be fairly small. Scharfstein and Robins (2002) and Rotnitzky et al. (2007) develop a sensitivity analysis that can be used to see how sensitive an estimator that assumes sequential ignorability of censoring is to differing amounts of residual dependence. Unfortunately, this method cannot be used on IPCW estimators as they used a different estimator that assumes sequential ignorability of censoring.

Despite this, the sensitivity analysis from Rotnitzky et al. (2007) is still applied to the Liver Registration data set. It is found that the bounds on the estimator that are derived are too wide to be of use practically. The method is also so computationally intensive that it is not easy to include many covariates or factors with many levels.

The other sensitivity analyses presented in Chapter 3 assess the sensitivity of the results from standard methods to the assumption of informative censoring. Sensitivity analyses for both parametric survival models and the Cox proportional hazards model are included.

The sensitivity analyses for parametric survival models are computationally simpler but cannot be applied to every data set as they require the marginal distributions of T and C to be one of the standard parametric survival distributions, such as the exponential of the Weibull. The sensitivity analysis in Siannis (2004) and Siannis et al. (2005) is our preferred sensitivity analysis for parametric survival models as it gives values that seem more feasible than the sensitivity analysis in Zhang and Heitjan (2006). However, the sensitivity analysis in Siannis (2004) and Siannis et al. (2005) does use several simplifying approximations which may affect the accuracy of the method.

The sensitivity analysis in Siannis (2011) uses a similar assumption about the dependence between T and C as Siannis (2004) and Siannis et al. (2005) and some of the same simplifying approximations but for the Cox proportional hazards model. It is more com-

putationally intensive as it requires the estimation of the baseline hazard functions but can be applied to a greater number of data sets.

The sensitivity analysis in Huang and Zhang (2008) is also for the Cox proportional hazards model but uses the same assumption as Zheng and Klein (1994), where the joint distribution of T and C is specified using a copula function. This is much more computationally intensive than Siannis (2011) and also requires additional untestable assumptions. This is because we have to specify the copula family to be used as well as the amount of dependence between T and C. It is for these reasons that the sensitivity analysis in Siannis (2011) is our preferred sensitivity analysis for the Cox proportional hazards model.

All of the estimators and methods described in the literature review in Chapters 2 and 3 rely on untestable assumptions to make the joint distribution of T and C identifiable. This means that we are unable to say which of the methods has the most realistic model for the liver transplantation setting. So any recommendations about which methods to use when analysing data are based on the properties of the methods and the intuitiveness of the assumption made about the dependence between T and C.

There are two main conclusions that can be drawn from the literature review: sensitivity analyses are useful for assessing the sensitivity of standard results to the assumption of informative censoring and IPCW estimators are the preferred estimators when carrying out analyses on a data set where we know the standard methods are sensitive to informative censoring. These conclusions have influenced the work in Chapters 4 and 6. In Chapter 4, we develop a new sensitivity analysis that overcomes some of the weaknesses of the sensitivity analyses discussed in Chapter 3. In Chapter 6, we use weights similar to those used for IPCW estimators to adjust for the informative censoring in the data set.

As discussed in Chapter 3, when applying a sensitivity analysis to the data set, we have to choose between using parametric models or Cox proportional hazards models for the marginal distributions of T and C. Using parametric models allows us to use a sensitivity that is simpler to apply but these models are not suitable for all data sets. Conversely, proportional hazards models are more flexible and so can be used for a wider range of data sets but the sensitivity analysis that has to be used is more computationally intensive. The new sensitivity methodology that we derive in Chapter 4 is a compromise between the two types of sensitivity analysis considered previously. We use piecewise exponential models for the marginal distributions of both T and C, which are more flexible than standard parametric survival models but allow us to retain the computationally simplicity of the sensitivity analysis.

There is only one drawback to using piecewise exponential survival models for the marginal distributions of T and C. To specify the distribution, suitable cut points for the intervals need to be specified. However, there is no preferred method for doing this in the

literature.

The sensitivity analysis in Chapter 4 is derived first for scalar parameters in each interval and is then extended to include covariates as well. There are two possible ways of incorporating covariates into the sensitivity analysis, using either a linear predictor or considering a vector of parameters. The sensitivity analysis for a linear predictor is simpler but the sensitivity analysis for a vector of parameters is more useful as the change in individual parameter estimates can be assessed. These two methods also give very different values of the estimated changes in parameter estimates, therefore the model that accounts for informative censoring is fitted to the Liver Registration data set to assess which is the more accurate method. The sensitivity analysis is used to approximate the parameter estimates for this model as it is time consuming to fit this model. It is found that the results from the sensitivity analysis for the vector of parameters are closest to those from this fitted model. Therefore this is our preferred method of incorporating covariates into the sensitivity analysis.

Another issue with this sensitivity analysis is that only small values of the parameter specifying the dependence between T and C can be used due to the approximations that are required to obtain the form of the sensitivity analysis equation. It is also useful to know how these approximations affect the accuracy of the sensitivity analysis. This is why a simulation study is conducted in Chapter 5. The simulation study uses a range of different parameter combinations so that the general applicability of the sensitivity analysis can be assessed. For simplicity, only models with scalar parameters in each interval are considered in the simulation study. Also, we only assess the accuracy of the sensitivity analysis when the piecewise parametric models are correctly specified. It is found that the sensitivity analysis tends to overestimate the change in the parameter estimates, but it is least accurate when there is a large amount of censoring in the data set or any individuals with particularly large observation times. Both of these are observed in the Liver Registration data set, so the sensitivity analysis should be more accurate in other applications than it is for the situation under consideration in this thesis. As expected, the sensitivity analysis also becomes less accurate as the value of the dependence parameter is increased.

The results of the simulation study are used in an attempt to improve the accuracy of the sensitivity analysis derived in Chapter 4. It is possible that including more terms in some of the approximations may improve the accuracy of the sensitivity analysis so a separate sensitivity analysis that uses a quadratic term in one of its Taylor expansions is also developed. However, it is found that for the Liver Registration data set this sensitivity analysis is not more accurate than the original sensitivity analysis.

Finally, in Chapter 6, a method that is particularly useful to NHSBT is considered.

The sequential stratification method allows the survival benefit of groups of patients on the waiting list for a liver transplant to be calculated. This is achieved by comparing the survival of each transplanted patient in the data set with the survival of suitable control patients. We amend an existing method so that the method used to match control patients is more realistic and the method is suitable for UK data rather than the US data for which it had originally been designed. The original method matched control patients by the length of time spent on the waiting list and we revised the method so that patients who were on the waiting list on the date of a transplant were used as control patients. However, it is found that the results for the Liver Registration data set are robust to the method of matching control patients used. We also consider using only patients that are blood group compatible with the donor organ as control patients and recommend using this criterion when applying the sequential stratification method to UK data. An additional criterion that can be used is to ensure that control patients have a suitable weight relative to the weight of the donor. However, it was found that this made the groups of control patients for each transplant too small so we recommend that this is not used.

The sequential stratification method can also be used to calculate the survival benefit of alternative transplantation therapies, such a split liver or a liver from an extended criteria donor. However, we only discuss this briefly and do not apply this to the Liver Registration data set.

As the survival benefit of the groups of patients on the waiting list is found using observational data, then we need to be aware that there may be bias in the results as a consequence of this. One particular example is selection bias. Any patients in the data set who were transplanted had been selected by clinicians as suitable for transplantation. Therefore, the results for each group of patients may not be applicable for every patient in that group.

7.2 Summary of strengths and weaknesses

There has not previously been a comprehensive review of the most recent literature on informative censoring, which we have carried out in Chapter 3. This is useful even if we are considering only the liver transplant setting as many of the findings also apply more general settings.

The main strength of the sensitivity analysis method derived in Chapter 4 is that it can be applied to a wide range of datasets whilst still being computationally simple due to the flexibility of piecewise exponential models. However, there are a couple of drawbacks for this method. The first is that it can only be applied for fairly small values of dependence between T and C due to the approximations necessary to derive the sensitivity analysis

equation. Secondly, we have to fit a piecewise exponetial model to the data and there is no preferred method for identifying the correct cut points to use.

Another strength of the thesis is that a simulation study has been carried out for the sensitivity analysis developed in Chapter 4 to assess its accuracy in a range of situations. This has not been done for the sensitivity analysis in Siannis et al. (2005) and Siannis (2004), which our method is based on. It is found that the sensitivity analysis using scalar parameters performed worse when there is a large amount of censoring or particularly large observation values, both of which are present in the liver transplantation setting. But the simulation study also demonstrates the general applicability of the sensitivity analysis as it was shown that it is fairly accurate for a wide range of parameter combinations.

The survival benefit methodology described in Chapter 6 has not been applied to UK data before. Some modifications of the method are also made to make it more suitable for this data. However, the drawback of this survival benefit methodology is that it is only suitable for the transplantation setting and cannot be applied in more general settings.

7.3 Suggestions for general data with potentially informative censoring

A flowchart summarising the process that we recommend should be followed if there is potentially informative censoring in a dataset is given in Figure 7.1.

We can see that in Figure 7.1, the first decision to be made is whether there is a convincing argument for potentially informative censoring in the data set. Unfortunately, due to the identifiability issues described in Section 1.1.2, it is not possible to develop a test to establish whether there is informative censoring in a data set. Therefore, the best alternative is to see whether there is a good argument for informative censoring and then conduct a sensitivity analysis to establish whether the assumption of informative censoring affects the results of the standard models.

When applying a sensitivity analysis to assess the sensitivity of the results from standard models to the assumption of informative censoring, we recommend using the sensitivity analysis we developed in Chapter 4. This is because it is flexible enough to be applied to most data sets whilst still being computationally simple. To establish whether the results of the standard models are sensitive to informative censoring for the application being considered, then the change in values of interest should be investigated. The values of interest that are used will depend on the application being considered. For example, in the liver transplantation setting the individual survival functions are assessed as ultimately we will be calculating survival benefit which is affected by changes in individual survival. It is not possible to develop a test of whether the changes in the values of interest

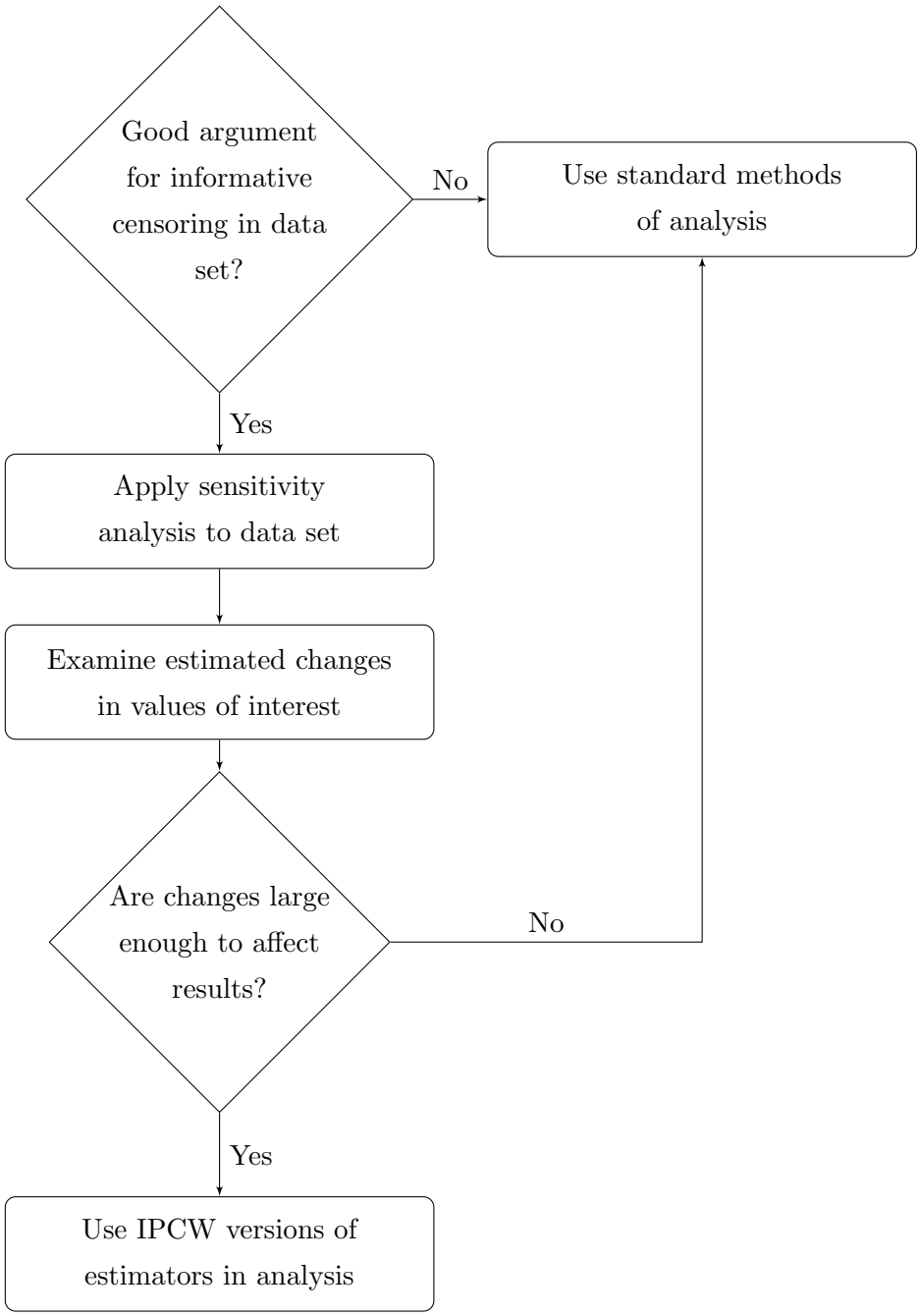

Figure 7.1: Flowchart showing the process to be followed if there is potentially informative censoring in a data set.

are significantly large. Therefore, the decision of whether the changes are considered to be large will again depend on the particular application that is being considered and is subjective.

If it is determined using the sensitivity analysis that the change in the values of interest are indeed considered large, then the analyses for this data should be carried out using IPCW versions of the required estimators. If the changes in the values of interest are found to be fairly small, then the standard methods of analysis can be used.

7.4 Summary of value of work to NHSBT

Chapters 2 and 3 provide a comprehensive review of the informative censoring methods in the literature for the liver transplantation setting. We identify the most suitable estimators and sensitivity analyses to be used in this setting. The results of this literature review are also applicable to the analysis of patients on the waiting list for transplants for other organs, with the exception of those waiting for a kidney transplant.

In Chapters 4 and 5, we develop an improved sensitivity analysis and establish its general applicability. Its flexibility and computational simplicity mean that it can be easily applied to any data set in the transplantation setting where there is potentially informative censoring.

The survival benefit methodology derived in Chapter 6 is particularly useful for NHSBT as it allows the survival benefit of groups of patients on the waiting list to be calculated. Modifications were made to the method presented in Schaubel et al. (2009b) to ensure that it is suitable for UK data, rather than the US data for which it was originally developed. It can also be easily amended to give the survival benefit of patients that receive new or alternative transplant therapies. This is useful to NHSBT as they have introduced the use of split livers and extended criteria donor to increase the number of donor livers available.

7.5 Extensions and Future Work

When developing the sensitivity analysis in Chapter 4, only piecewise exponential models are considered. One possible extension of this work is to make it suitable for use with piecewise Weibull models.

The simulation study for this sensitivity analysis that is carried out in Chapter 5 could also be extended. We only consider the accuracy of the sensitivity analysis when the piecewise parametric models for the marginal distributions are correctly specified. However, it is likely that we will not identify the exact piecewise parametric distribution present in a data set, so it would be useful to investigate the robustness of the sensitivity

analysis to the misspecification of the models for the marginal distributions.

There is also more work that could be done on the survival benefit methodology derived in Chapter 6. The method could be applied to the transplantation of other organs as it is likely that there will be the same issues with informative censoring. We could also look into calculating the survival benefit of alternative therapies, mentioned briefly in Section 6.3.1, in more detail and develop programs to implement this.

Finally, we observe a large amount of missing data in our data set so another area of possible future work is to develop an multiple imputation method for large medical databases. This would prevent us from having to disregard large numbers of observations due to missing data.

Bibliography

[1] M. Aitken, D. Anderson, B. Francis, and J. Hinde. *Statistical Modelling in GLIM*. Clarendon Press, Oxford, 1989.

[2] M. Aitken, N. Laird, and B. Francis. A reanalysis of the Stanford heart transplant data. *Journal of the American Statistical Society*, 78:264–274, 1983.

[3] P. D. Allison. *Survival Analysis Using the SAS System: A Practical Guide*. Cary, NC: SAS Institute Inc., 1995.

[4] S. K. Asrani, W. R. Kim, and J. K. Heimbach. Survival benefit of liver transplantation: One size fits all or fits none? *Hepatology*, 50:352–354, 2009.

[5] A. W. Avolio, S. Agnes, A. Gasbarrini, E. Nure, M. Siciliano, and M. Castagneto. Prognostic value of meld score and donor quality in liver transplantation: Implications for the donor recipient match. *Transplantation Proceedings*, 38:1059–1062, 2006.

[6] K. Barber, S. Madden, J. Allen, D. Collett, J. Neuberger, and A. E. Gimson. Elective liver transplant list mortality; development of a united kingdom end-stage liver disease (ukeld) score. *Transplantation*, 92:469–476, 2011.

[7] K. M. Barber, S. Pioli, and J. E. Blackwell. Development of a UK score for patients with end-stage liver disease. *Hepatology*, 46:510A, 2007.

[8] P. Bickel, C. Klassen, Y. Ritov, and J. Wellner. *Efficient and Adaptive Inferences in Semiparametric Models*. Springer, 1998.

[9] R. Braekers and N. Veraverbeke. The partial Koziol-Green model with covariates. *Journal of Statistical Planning and Inference*, 92:55–71, 2001.

[10] R. Braekers and N. Veraverbeke. Testing for the partial Koziol-Green model with covariates. *Journal of Statistical Planning and Inference*, 115:181–192, 2003.

[11] R. Braekers and N. Veraverbeke. Cox's regression model under partially informative censoring. *Communications in Statistics - Theory and Methods*, 34:1793–1811, 2005.

[12] R. Braekers and N. Veraverbeke. A conditional Koziol-Green model under dependent censoring. *Statistics & Probability Letters*, 78:927–937, 2008.

[13] M. Brendel. Testing hypotheses under a generalized Koziol-Green model with partially informative censoring. *Statistics*, 39:329–345, 2005.

[14] N. Breslow. Covariance analysis of censored survival data. *Biometrics*, 30:89–99, 1974.

[15] E. Buggins. Allocations of organs to non-UK EU residents. Report, available from Department of Health website: *http://www.dh.gov.uk/en/Publicationsandstatistics /Publications/PublicationsPolicyAndGuidance/DH_103515*, 2009.

[16] A. B. Cantor. *SAS Survival Analysis Techniques for Medical Research*. Cary, NC: SAS Institute Inc, second edition, 2003.

[17] D. G. Clayton. A model for association in bivariate lifetables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, 65:141–151, 1978.

[18] D. Collett. *Modelling Survival Data in Medical Research*. Chapman & Hall/CRC, second edition, 2003.

[19] D. R. Cox. The analysis of exponentially distributed life-times with two types of failure. *Journal of the Royal Statistical Society Series B*, 21:411–421, 1959.

[20] D. R. Cox. Regression models and life tables (with discussion). *Journal of the Royal Statistical Society Series B*, 74:187–220, 1972.

[21] M. Crowder. On the identifiability crisis in competing risks analysis. *Scandinavian Journal of Statistics*, 18:223–233, 1991.

[22] M. Crowder. Identifiability crises in competing risks. *International Statistical Review*, 62:379–391, 1994.

[23] M. Crowder. *Classical Competing Risks*. Chapman & Hall/CRC, 2001.

[24] V. DeGruttola and X. M. Tu. Modeling the progression of CD4-Lymphocyte Count and its relationship to survival time. *Biometrics*, 50:1003–1014, 1994.

[25] B. Efron. The two sample problem with censored data. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 4*, pages 831–853. University of California Press, Berkeley, California, 1976.

[26] L. Fisher and P. Kanarek. Presenting censored survival data when censoring and survival times may not be independent. In F. Proschan and R. Serfling, editors, *Reliability and Biometry: Statistical Analysis of Lifelength*, pages 303–326. Philadelphia: SIAM, 1974.

[27] T. R. Fleming and D. P. Harrington. *Counting Processes and Survival Analysis*. New York: Wiley, 1991.

[28] R. B. Freeman, N. Jamieson, D. E. Schaubel, R. J. Porte, and F. G. Villamil. Who should get a liver graft? *Journal of Hepatology*, 50:664–673, 2009.

[29] M. Friedman. Piecewise exponential models for survival data with covariates. *Annals of Statistics*, 10:101–113, 1982.

[30] U. Gather and J. Pawlitschko. Estimating the survival function under a generalized koziol-green model with partially informative censoring. *Metrika*, 48:189–207, 1998.

[31] J. W. Hogan and N. M Laird. Model-based approaches to analysing incomplete longitudinal and failure time data. *Statistics in Medicine*, 16:259–272, 1997.

[32] F. Hu and J. V. Zidek. The weighted likelihood. *Canadian Journal of Statistics*, 30:347–371, 2002.

[33] X. Huang and R. A. Wolfe. A frailty model for informative censoring. *Biometrics*, 58:510–520, 2002.

[34] X. Huang and N. Zhang. Regression survival analysis with an assumed copula for dependent censoring: A sensitivity analysis approach. *Biometrics*, 64:1090–1099, 2008.

[35] M. Jacob, L. P. Copley, J. D. Lewsey, A. Gimson, M. Rela, and J. H. P. van der Meulen. Functional status of patients before liver transplantation as a predictor of posttransplant mortality. *Transplantation*, 80:52–57, 2005.

[36] J. D. Kalbfleisch and R. L. Prentice. *The Statistical Analysis of Failure Time Data*. Wiley, 2002.

[37] W. R. Kim and W. K. Kremers. Benefits of "the benefit model" in liver transplantation. *Hepatology*, 48:697–698, 2008.

[38] J. P. Klein and M. L. Moeschberger. Bounds on net survival probabilities for dependent competing risks. *Biometrics*, 44:529–538, 1988.

[39] J. A. Koziol and S. B. Green. A Cramér-von Mises statistic for randomly censored data. *Biometrika*, 63:465–474, 1976.

[40] S. W. Lagakos. General right censoring and its impact on the analysis of survival data. *Biometrics*, 35:139–156, 1979.

[41] W. A. Link. A model for informative censoring. *Journal of the American Statistical Association*, 84:749–752, 1989.

[42] R. J. A. Little. Modeling the drop-out mechanism in repeated-measure studies. *Journal of the American Statistical Association*, 90:1112–1121, 1995.

[43] M. R. Lucey, D. E. Schaubel, M. K. Guidinger, S. Tome, and R. M. Merion. Effect of alcoholic liver disease and hepatitis c infection on waiting list and posttransplant mortality and transplant survival benefit. *Hepatology*, 50:400–406, 2009.

[44] D. G. Maluf, E. B. Edwards, and H. M. Kauffman. Utilization of extended donor criteria liver allograft: Is the elevated risk of failure independent of the model for end-stage liver disease score of the recipient. *Transplantation*, 82:1653–1657, 2006.

[45] R. M. Merion, D. E. Schaubel, D. M. Dykstra, R. B. Freeman, F. K. Port, and R. A. Wolfe. The survival benefit of liver transplantation. *American Journal of Transplantation*, 5:307–313, 2005.

[46] R. B. Nelsen. *An Introduction to Copulas*. Springer-Verlag New York, Inc., 1999.

[47] J. Neuberger, A. Gimson, M. Davies, M. Akyol, J. O'Grady, A Burroughs, and M. Hudson. Selection of patients for liver transplantation and allocation of donated livers in the UK. *Gut*, 57:252–257, 2008.

[48] W. K. Newey. Semiparametric efficiency bounds. *Journal of Applied Econometrics*, 5:99–135, 1990.

[49] D. Oakes. A model for association in bivariate survival data. *Journal of the Royal Statistical Society, Series B*, 44:412–422, 1982.

[50] S. J. Pelletier, D. E. Schaubel, G. Wei, M. J. Englesbe, J. D. Punch, R. A. Wolfe, F. K. Port, and R. M. Merion. Effect of body mass index on the survival benefit of liver transplantation. *Liver Transplantation*, 13:1678–1683, 2007.

[51] A. V. Peterson. Bounds for a joint distribution function with fixed sub-distribution functions: application to competing risks. *Proceedings of the National Academy of Sciences*, 73:11–13, 1976.

[52] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, second edition, 1992.

[53] L. Rivest and M. T. Wells. A martingale approach to the copula-graphic estimator for the survival function under dependent censoring. *Journal of Multivariate Analysis*, 79:138–155, 2001.

[54] M. S. Roberts, D. C. Angus, C. L. Bryce, Z. Valenta, and L. Weissfeld. Survival after liver transplantation in the united states: A disease-specific analysis of the unos database. *Liver Transplantation*, 10:886–897, 2004.

[55] J. M. Robins. Information recovery and bias adjustment in proportional hazards regression analysis of randomized trials using surrogate markers. In *Proceedings of the Biopharmaceutical Section, American Statistical Association*, pages 24–33. Alexandria, VA: American Statistical Association, 1993a.

[56] J. M. Robins. Analytic methods for estimating HIV-treatment and cofactor effects. In D. G. Ostrow and R. Kessler, editors, *Methodological Issues of AIDS Behavioral Research*, pages 213–287. New York: Plenum Publishing, 1993b.

[57] J. M. Robins and D. H. Finkelstein. Correcting for non-compliance and dependent censoring in an aids clinical trial with inverse probability of censoring weighted (IPCW) log-rank tests. *Biometrics*, 56:779–788, 2000.

[58] J. M. Robins and A. Rotnitzky. Recovery of information and adjustment for dependent censoring using surrogate markers. In N. Jewell and K. Dietz, editors, *AIDS Epidemiology: Methodological Issues*, pages 297–331. Boston: Birkhauser, 1992.

[59] A. Rotnitzky, A. Bergesio, and A. Farall. Analysis of quality-of-life adjusted failure time data in the presence of competing, possibly informative, censoring mechanisms. *Lifetime Data Analysis*, 15:1–23, 2009.

[60] A. Rotnitzky, A. Farall, A. Bergesio, and D. Scharfstein. Analysis of failure time data under competing censoring mechanisms. *Journal of the Royal Statistical Society Series B*, 69:307–327, 2007.

[61] P. K. Ruan and R. J. Gray. Sensitivity analysis of progression-free survival with dependent withdrawal. *Statistics in Medicine*, 27:1180–1198, 2008.

[62] G. A. Satten, S. Datta, and J. Robins. Estimating the marginal survival function in the presence of time dependent covariates. *Statistics & Probability Letters*, 54:397–403, 2001.

[63] D. O. Scharfstein and J. M. Robins. Estimation of the failure time distribution in the presence of informative censoring. *Biometrika*, 89:617–634, 2002.

[64] D. O. Scharfstein, J. M. Robins, W. Eddings, and A. Rotnitzky. Inference in randomized studies with informative censoring and discrete time-to-event endpoints. *Biometrics*, 57:404–413, 2001.

[65] D. E. Schaubel, M. K. Guidinger, S. W. Biggins, J. D. Kalbfleisch, E. A. Pomfret, P. Sharma, and R. M. Merion. Survival benefit-based deceased-donor liver allocation. *American Journal of Transplantation*, 9:970–981, 2009a.

[66] D. E. Schaubel, C. S. Sima, N. P. Goodrich, S. Feng, and R. M. Merion. The survival benefit of deceased donor liver transplantation as a function of candidate disease severity and donor quality. *American Journal of Transplantation*, 8:419–425, 2008.

[67] D. E. Schaubel, R. A. Wolfe, and F. K. Port. A sequential stratification method for estimating the effect of a time-dependent treatment in observational studies. *Biometrics*, 62:910–917, 2006.

[68] D. E. Schaubel, R. A. Wolfe, C. S. Sima, and R. M. Merion. Estimating the effect of a time-dependent treatment by levels of an internal time-dependent covariate: Application to the contrast between liver wait-list and posttransplant mortality. *Journal of the American Statistical Association*, 104:49–59, 2009b.

[69] M. D. Schluchter. Methods for the analysis of informatively censored longitudinal data. *Statistics in Medicine*, 11:1861–1870, 1992.

[70] P. Sharma, D. E. Schaubel, M. K. Guidinger, and R. M. Merion. Effect of pretransplant serum creatinine on the survival benefit of liver transplantation. *Liver Transplantation*, 15:1808–1813, 2009.

[71] F. Siannis. Applications of a parametric model for informative censoring. *Biometrics*, 60:704–714, 2004.

[72] F. Siannis. Sensitivity analysis for multiple right censoring processes: Investigating mortality in psoriatic arthritis. *Statistics in Medicine*, 30:356–367, 2011.

[73] F. Siannis, J. Copas, and G. Lu. Sensitivity analysis for informative censoring in parametric survival models. *Biostatistics*, 6:77–91, 2005.

[74] E. V. Slud and L. V. Rubinstein. Dependent competing risks and summary survival curves. *Biometrika*, 70:643–649, 1983.

[75] P. E. Spector. *SAS Programming for Researchers and Social Scientists*. SAGE Publications, 1993.

[76] T. M. Therneau and P. M. Grambsch. *Modelling Survival Data: Extending the Cox Model*. New York: Springer, 2000.

[77] A. B. Troxel, G. Ma, and D. F. Heitjan. An index of local sensitivity to nonignorability. *Statistica Sinica*, 14:1221–1237, 2004.

[78] W. Tsai and J. Crowley. A large sample study of the generalised maximum likelihood estimators from incomplete data via self-consistency. *Annals of Statistics*, 13:1317–1334, 1985.

[79] A. A. Tsiatis. A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences*, 72:20–22, 1975.

[80] J. H. P. van der Meulen, J. D. Lewsey, M. F. Dawwas, and L. P. Copley. Adult orthotopic liver transplantation in the United Kingdom and Ireland between 1994 and 2005. *Transplantation*, 84:572–579, 2007.

[81] M. L. Volk, A. S. F. Lok, S. J. Pelletier, P. A. Ubel, and R. A. Hayward. Impact of the model for end-stage liver disease allocation policy on the use of high-risk organs for liver transplantation. *Gastroenterology*, 135:1568–1574, 2008.

[82] R. H. Weisner, S. V. McDiarmid, and P. S. Kamath. MELD and PELD: Application of survival models to liver allocation. *Liver Transplantation*, 7:567–580, 2001.

[83] M. C. Wu and K. R. Bailey. Estimation and comparison of changes in presence of informative right censoring: Conditional linear model. *Biometrics*, 45:939–955, 1989.

[84] M. C. Wu and R. J. Carroll. Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process. *Biometrics*, 44:175–188, 1988.

[85] J. Yan. Enjoy the joy of copulas: With a package copula. *Journal of Statistical Software*, 21:1–21, 2007.

[86] D. Zeng. Estimating marginal survivor function by adjusting for dependent censoring using many covariates. *Annals of Statistics*, 32:1533–1555, 2004.

[87] H. Zhang and M. B. Rao. On generalised maximum likelihood estimation in the proportional hazards model with partially informative censoring. *Metrika*, 59:125–136, 2004.

[88] J. Zhang and D. F. Heitjan. A simple local sensitivity analysis tool for nonignorable coarsening: Application to dependent censoring. *Biometrics*, 62:1260–1268, 2006.

[89] M. Zhang and D. E. Schaubel. Estimating differences in restricted mean lifetime using observational data subject to dependent censoring. *Biometrics*, 67:740–749, 2011.

[90] M. Zheng. *On the Use of Copulas in Dependent Competing Risks*. PhD thesis, Ohio State University, 1992.

[91] M. Zheng and J. P. Klein. A self-consistent estimator of the marginal survival functions based on dependent competing risk data and an assumed copula. *Communications in Statistics A*, 23:2299–2311, 1994.

[92] M. Zheng and J. P. Klein. Estimates of marginal survival for dependent competing risks based on an assumed copula. *Biometrika*, 82:127–138, 1995.

[93] M. Zhou. Understanding the Cox regression models with time-change covariates. *American Statistician*, 55:153–155, 2001.