2011
2012

FACULTY OF SCIENCES

Master of Statistics: Biostatistics

Masterproef

Correcting for cross-over bias in randomized controlled clinical trials

Promotor :
Prof. dr. Ziv SHKEDY

Promotor :
Dr. SUZY VAN SANDEN

Agnes Natukunda

Master Thesis nominated to obtain the degree of Master of Statistics , specialization Biostatistics

De transnationale Universiteit Limburg is een uniek samenwerkingsverband van twee universiteiten in twee landen: de Universiteit Hasselt en Maastricht University

universiteit hasselt
KNOWLEDGE IN ACTION

Maastricht University

Universiteit Hasselt | Campus Diepenbeek | Agoralaan Gebouw D | BE-3590 Diepenbeek
Universiteit Hasselt | Campus Hasselt | Martelarenlaan 42 | BE-3500 Hasselt

Maastricht University

universiteit hasselt
KNOWLEDGE IN ACTION

2011
2012

FACULTY OF SCIENCES

Master of Statistics: Biostatistics

Masterproef

Correcting for cross-over bias in randomized controlled clinical trials

Promotor :
Prof. dr. Ziv SHKEDY

Promotor :
Dr. SUZY VAN SANDEN

Agnes Natukunda

Master Thesis nominated to obtain the degree of Master of Statistics , specialization Biostatistics

Maastricht University

universiteit hasselt
KNOWLEDGE IN ACTION

To my mum Rosemary Bitakashoborokire and my late dad Paul Bitakashoborokire

Acknowledgment

I give thanks to God the Almighty for the gift of life and energy that He has bestowed upon me and to His guidance not only during the time of this project but in my entire life.

I take this opportunity to deeply express my gratitude to my supervisors Prof. dr. Ziv SHKEDY and Dr. Suzy VAN SANDEN for their wonderful guidance and pieces of advice they have offered me in the accomplishment of this work.

I would like to deeply express my appreciation to Hasselt University professors for the wealth of knowledge they have imparted in me for the two years of my master programme in Biostatistics.

Many thanks go to my mum, brothers and sisters for their encouragements, prayers, guidance, moral support, advices and all sorts of support they have offered me in my entire life; God bless you abundantly . To my friends, classmates, group-members thank you very much for your encouragements and support at those difficult and challenging moments.

My last acknowledgement goes to VLIR-UOD (Flemish University Development Cooperation) for financing my master studies in biostatistics. I will always live to appreciate and be thankful of your support.

Contents

List of Tables

List of Figures

List of abbreviations

AML	Acute Myeloid Leukemia
ECOG	Cooperative Oncology Group
IPCW	Inverse Probability of Censoring Weights
ITT	Intention –To-Treat
OS	Overall Survival
PFS	Progression Free Survival
RCTs	Randomised Controlled Trials
RPSFTM	Rank Preserved Structural Failure Time Models
SNM	Structural Nested Models

Abstract

Ideally, therapeutic interventions are evaluated through randomized clinical trials. These trials are commonly analysed with an intent-to-treat (ITT) approach, whereby patients are analyzed in their assigned treatment group regardless of actual treatment received. If an interim analysis of such trials demonstrates compelling evidence of a difference in benefit, ethical considerations often dictate that the trial be unblinded and participants be provided access to the more efficacious agent. Because interim analysis may not address longer-term outcomes of interest, important clinical questions such as overall survival benefit—the ultimate test of efficacy to many—may remain unanswered. The ensuing crossover disturbs randomization and may lead to biased longer-term analysis, compromising the utility of clinical data. In this thesis, I discuss the biases associated with ITT analysis and, alternatively, censoring of follow-up data after selective crossover. Moreover, discussed also is how the statistical procedure of inverse probability of censoring weighted (IPCW) analysis is used to account for selective crossover as an alternative to ITT or censoring analysis. The results showed that IPCW analysis was particularly suited for detecting overall survival benefits.

1 Introduction

Randomized controlled trials (RCTs) are widely used to assess the effects of a new treatment or procedure compared to a control treatment. Survival outcomes are commonly used in RCTs with the time to an event such as death or disease progression analysed. Advanced disease trials are often designed with progression free survival as the primary endpoint, and overall survival as a secondary endpoint (Morden, Lambert, Latimer, Abrams, & Wailoo, 2011).

In the conduct of randomized controlled trials, it is common for patients to switch from the treatment to which they were randomized to another trial treatment or to a non-trial treatment or stop receiving treatment altogether. Trial protocols often attempt to control these switches while maintaining a degree of flexibility over the treatment a patient can receive, although this varies greatly between trials, and switching remains common. Switches may occur for a number of reasons, many of which are related to an individual's prognosis (Morden, Lambert, Latimer, Abrams, & Wailoo, 2011).

One of the main difficulties in estimating overall survival is the confounding caused by treatment crossover. It is common to allow control group patients to switch on to the experimental treatment following disease progression for ethical reasons. Thus the overall survival advantage associated with the experimental treatment cannot be estimated with confidence based on the intention to treat (ITT) data, because a proportion of the patients randomised to the control treatment will have received the experimental treatment and vice versa. Equally a per protocol approach where patients who switch treatments are censored from the analysis at the time of switching-is also likely to be confounded because treatment switching is unlikely to have occurred at random. In such circumstances censoring is informative and the randomisation of the trial is compromised.

Treatment crossover has been an important issue in the analysis of clinical trials but the methods used to account for the impact of crossover on the treatment effect have generally been simplistic. Most regularly censoring approaches have been used but often the crossover has been ignored and standard ITT analyses conducted which produce heavily biased results. More recently statistical techniques to address the crossover problem have been devised, for example, Inverse probability of censoring weights (IPCW) and rank preserved structural failure time models (RPSFTM).

1.1 Objectives of the analysis

The primary objective of this analysis is to compare overall survival (OS) in patients 65 years and older who had newly diagnosed de novo or secondary AML and either poor or intermediate-risk cytogenetic who were randomly assigned to receive experimental treatment and control treatment. The focus will be on adjusting survival estimates due to treatment crossover by applying the Inverse probability of censoring weights (IPCW) method. I also demonstrate the differences in estimates obtained from using methods not accounting for crossover such as intent to treat and censoring crossovers with using a more complex method-Inverse probability of censoring weights (IPCW). In addition, comparison between treatment arms will be done for time to disease progression.

The rest of thesis is organized as follows. In chapter 2, some back ground of the clinical trial and the description of the data are presented whereas chapter 3 discusses the methodology applied in the analysis starting with methods not accounting for crossover and then the IPCW method. In chapter 4 I summarize the experience of using the ITT and Censored methods versus the IPCW method by presenting the results from the fitted models. At last, Chapter 5 provides a summary of the discussion on this study and further analysis methods on the subject are discussed.

2 Data description

2.1 Study population

Data on disease progression and death were collected in a randomized , open label, phase III study in patients with newly diagnosed, histologically confirmed de novo or secondary AML to compare treatments A and B which in this thesis will be referred to as experimental treatment and control treatment respectively. Randomisation was stratified by Eastern Cooperative Oncology Group (ECOG) performance status (0-1 versus 2), age (65-69 versus 70+ years), and cytogenetic risk (poor versus intermediate). Patients could continue in the study until disease progression or death or if they experienced unacceptable toxicity or inter-current illness that prevented further administration of the treatment. For this analysis both time to death and time to disease progression will be used to compare the experimental treatment against the control treatment.

2.2 Description of variables

The variables and coding considered in this analysis are presented in table 1. The data are right censored and most of the covariates are binary apart from age which is a continuous variable.

Table 1: List of variables in the data set

Variables	*Description*	*Codes/values*
AMLtype	Acute Myeloid Leukemia	1=de novo, 2=Secondary
ECOG	Cooperative Oncology Group performance status	1=0-1, 2=2
AGE	Age	Years
BMBLAST	Bone Marrow Blasts (% of abnormal (leukemia) cells)	1="<= 30%", 2=">30%"
RACE	Race	1=White, 2=Non-white
GENDER	Gender	1=Male, 2=Female
CGRISK	Cytogenetic risk, based on chromosomal abnormalities	1=Poor, 2=Intermediate
TRTMNT	2 Treatment arms	0=Control, 1=treatment
TIME_TO_DEATH		Days
STATUS_DEATH	Censoring status for Death	0=Alive, 1=Dead
TIME_TO_PFS		Days
STATUS_PFS	Censoring status for PFS	0=No progression, 1=Progression

2.2.1 Patient characteristic and risk factors by outcome

First, descriptive statistics were used to give an insight into the distribution of the variables. Summaries of the baseline characteristics of patients in the two treatment groups are shown in table 2. There were no missing data for all the covariates as well as outcomes.

A total of 485 patients were enrolled in the trial with almost equal number of patients in each treatment arm. Clinical characteristics were well balanced across the treatment groups as shown in table 3. The median follow-up period of 185 days after randomisation was observed.

Table 2: Patient characteristics

variable	*Treatment n(%)*	*Control n(%)*
No of patients	242	243
Median Age, years(Q1;Q3)	73(69;77)	73(69;78)
Age range(years)	65-89	65-91
Gender		
Male	137(56.61)	151(62.14)
Female	105(43.39)	92(37.86)
ECOG status		
0-1	182(75.21)	178(73.25)
2	60(24.79)	65(26.75)
Cgrisk		
Poor	88(36.36)	87(35.8)
Intermediate	154(63.64)	156(64.2)
Race		
White	209(86.36)	213(87.65)
Non-white	33(13.64)	30(12.35)
AMLtype		
De novo	155(64.05)	157(64.61)
Secondary	87(35.95)	86(35.39)
Bmblast		
<= 30%	70(28.93)	68(27.98)
>30%	172(71.07)	175(72.02)

Table 3: Risk factor by outcome

Table 2: Risk factor by outcome

	Outcome							
	Time to death_ITT		*Time to death_CEN*		*Time to PFS_ITT*		*Time to PFS_CEN*	
Risk factor	*Dead n(%)*	*Alive n(%)*	*Dead n(%)*	*Alive n(%)*	*Progression n(%)*	*No progression n(%)*	*Progression n(%)*	*No progression n(%)*
Gender								
Male	267(59.87)	21(53.85)	228(61.13)	60(53.57)	278(59.91)	10(47.62)	268(60.63)	20(46.51)
Female	179(40.13)	18(46.15)	145(38.87)	52(46.43)	186(40.09)	11(52.38)	174(39.37)	23(53.49)
Ecog status								
0-1	328(73.54)	32(82.05)	269(72.12)	91(81.25)	344(74.14)	16(76.19)	323(73.08)	37(86.05)
2	118(26.46)	7(17.95)	104(27.88)	21(18.75)	120(25.86)	5(23.81)	119(26.92)	6(13.95)
Cgrisk								
Poor	166(37.22)	9(23.08)	139(37.27)	36(32.14)	171(36.85)	4(19.05)	161(36.43)	14(32.56)
Intermediate	280(62.78)	30(76.92)	234(62.73)	76(67.86)	293(63.15)	17(80.95)	281(63.57)	29(67.44)
Race								
White	387(86.77)	35(89.74)	324(86.86)	98(87.5)	405(87.28)	17(80.95)	385(87.1)	37(86.05)
Non-white	59(13.23)	4(10.26)	49(13.14)	14(12.5)	59(12.72)	4(19.05)	57(12.9)	6(13.95)
AMLtype								
De novo	286(64.13)	26(66.67)	245(65.68)	67(59.82)	297(64.01)	15(71.43)	285(64.48)	27(62.79)
Secondary	160(35.87)	13(33.34)	128(34.32)	45(40.18)	167(36.00)	6(28.57)	157(35.52)	16(37.21)
Bmblast								
<=30%	126(28.25)	12(30.77)	110(29.49)	28(25)	135(29.09)	3(14.29)	129(29.19)	9(20.93)
>30%	320(71.75)	27(69.23)	263(70.51)	84(75)	329(70.91)	18(85.71)	313(70.81)	34(79.07)

2.3 Software used

Statistical analysis was conducted using SAS version 9.2 and R version 2.12. All statistical tests were done at 5% level of significance unless stated otherwise.

3 Methodology

3.1 Statistical Analysis

There are several methods that have been developed to adjust survival estimates in the presence of crossover. Methods not accounting for cross-over such as intention-to-treat (ITT) and per protocol (censoring at time of crossover) have often been used although more complex randomisation-based approaches to adjusting for treatment crossover such as Rank Preserving Structural Failure Time Models (RPSFTM) and Iterative Parameter Estimation algorithm have been developed.

Complex observational-based methods such as Structural Nested Models (SNM), Iterative Parameter Estimation and Marginal Structural Models (MSM) also exist. Even though these methods were developed for observational studies, they can be applied in a randomized controlled trial (RCT) framework (Latimer, et al.). For this analysis, methods like MSM will not be considered here due to lack of time dependent covariates.

IPCW analysis is a method proposed to cope with the issue of crossover. The dependence of censoring on disease outcome–associated factors is used to up or down weight participants remaining at risk, in essence filling in for their crossed over and censored fellow participants. In simpler terms, IPCW analysis attempts to model what results would be had no selective crossover occurred (Rimawi & Hilsenbeck, 2012). Thus we focus on the IPCW approach.

3.1.1 Methods not accounting for crossover

Methods not accounting for crossover are discussed before proceeding to IPCW method. These methods represent simple techniques that have often been used to analyse data in which treatment crossover has occurred. These methods were included in the analysis to make comparisons with IPCW method. Methods not accounting for crossover performed in this analysis include intention-to-treat and censored analysis (censor patients at time of

crossover). An analysis that excludes patients who switch treatment was not done here since there is a high percentage of patients who switch treatment.

3.1.2 ITT analysis

The intention-to-treat analysis is based on the principle that a patient's decision to stop or refuse treatment should be regarded as an integral part of treatment administration. Thus, if non-compliance in an RCT only manifests as patients stopping or refusing their randomized treatment, it can be argued that the intention-to-treat analysis addresses the clinically relevant question by comparing the overall consequences of treatments, including a patient's decision to stop. This argument, however, is difficult to sustain if patients actually switch from their allocated treatment to the alternative randomized treatment. In this situation, if one treatment is better than the other, then the intention-to-treat analysis will tend to underestimate this treatment effect. The extent of the bias created will depend on the size of any treatment effect, the proportion of patients who switch treatments and the duration of time they receive each therapy (Law & Kaldor, 1996).

An intention to treat analysis was done by fitting a multivariate cox proportional hazards regression model of the form;

$$h_i(t \mid x_i, Trt_i) = h_0(t)\exp(\beta x_i + \gamma Trt_i) \qquad \text{Model (1)}$$

Where: $h_i(t)$ is the hazard function for patient i

$h_0(t)$ is the baseline hazard function of a patient in the control group when all the covariates (x) are zero

Trt_i is the treatment (0=Control, 1=Experimental treatment)

β is the vector of regression parameters and $\exp(\beta)$ is the hazard ratio (HR)

γ is log hazard for the treatment

3.1.3 Censored analysis

When some patients switch treatments in an RCT, an alternative analysis which is sometimes performed with the goal of estimating the treatment effect more accurately is to censor the patients' survival time at the point of switching treatment. Although widely accepted, censored analysis has its limitations. For example, in studies in which selective crossover occurs, patients who elect to cross over may have different prognostic characteristics than those who do not crossover. Differences in prognostic characteristics were also observed in this trial, because 81.98% of those who switched treatment had ECOG status of 0-1 as compared to 72.64% who didn't switch treatment.

In theory, in a censored analysis, random cross over would tend to have no effect, preserving the true effects of the treatment. Problems arise when the crossover is not random, and censoring is informative (i.e. when censored patients are either more likely or less likely to experience specific events in future than are uncensored individuals). With biased crossover of patients with the worst prognoses, censoring removes them from further consideration, making the less effective therapy seem better, whereas biased crossover of patients with the best prognoses leaves those with the worst prognoses uncensored, making the less effective therapy seem worse. As suggested in the example, the realized crossover may be a complex mixture, the effects of which are virtually impossible to predict (Rimawi & Hilsenbeck, 2012).

The censored analysis was done by fitting a multivariate cox proportional hazard regression model (1). However, as several authors have noted, these analyses are prone to large biases (Law & Kaldor, 1996).

3.1.4 Inverse probability of censoring weights (IPCW)

The IPCW method adjusts for bias associated with time-dependent confounders that are affected by prior treatment or exposure (e.g. drop out due to adverse effects). However in

Randomized clinical trials, this method has also been applied to control for selective crossover (informative censoring), which is more likely among high risk subjects (Cain & Cole, 2009)

The weights are estimated to represent the inverse probability of informative censoring, given factors affecting the likelihood of crossover and/or survival. The Inverse probability of censoring weights method creates a scenario of missing follow-up data by censoring the follow-up of each subject at the time of cross over (i.e. they get weight equal to zero for time periods after cross over). Subjects with similar characteristics who did not cross over, in IPCW method they receive bigger weights in order to "re-create" the population that would have been observed in the absence of cross over (Latimer, et al.).

Specifically, at time point j, each participant is assigned a weight that is inversely proportional to the estimated conditional probability that the participant remained not artificially censored through time point j. The conditional probabilities and weights can be estimated by fitting a discrete time logistic regression model for artificial censoring, in which the common predictors of the endpoint of interest and the artificial censoring are included as covariates in the model (Toh, Hernández-Díaz, Logan, Robins, & Hernán, 2010).

The ability of IPCW to recapture unobserved survival data and yield an unbiased estimate had the artificial censoring mechanism never occurred is dependent on whether the assumptions of exchangeability and correct model specification are met (Howe, Cole, Chmiel, & Munoz, 2011). The exchangeability assumption implies that given the measured common predictors of the outcome of interest and artificial censoring, artificially censored participants have the same prognosis with respect to the outcome of interest as do participants who are not artificially censored.

Conducting an IPCW analysis for survival to adjust for cross over in randomized controlled trials is a three stage process. The first step entails creating a panel data. The original data file contained one record per patient but here a transformed data file with each patient having a record every 180 days was used. For all patients, follow-up time from randomization until failure (death) or informative censoring (crossover) was partitioned into fixed intervals (i.e., 6-months intervals) This file format was necessary to capture treatment history.

The second step is to estimate the probability to crossover given factors affecting likelihood of crossover and/or survival and then obtain weights for each observation (per time point) using a logistic regression models with crossover as the dependent variable. The model is of the form;

$$\log itP[C_{ij} = 0 \mid X_i] = \beta_0 + \beta X_i \qquad \text{Model (2)}$$

Where C_{ij} represents status of crossover for i^{th} patient at j^{th} time point with $C_{ij} = 0$ for patients who switched treatment whereas, X_i represents a vector of covariates for patient i and β is a vector of regression parameters. Thus the weights for each observation (per time point) were defined as the inverse of the probability to crossover. This means that the subject specific weights are time-varying. In this analysis there were various scenarios considered in estimating the probability to cross required to obtain weights and they are discussed below.

3.1.4.1 Weights based on baseline covariates only

The probability to crossover was taken as a function of only baseline covariates. This implies that equal probability to crossover regardless of the initial randomized treatment was assumed though we know this is not the case. The weights are inversely proportional to an estimate of the conditional probability of remaining uncensored until time point *j*. The logistic regression model for the discrete-time of hazard censoring was fitted, specifically;

$$\text{logit } P[C_{ij} = 0 \mid X_i] = \beta_0 + \beta X_i \qquad \text{Model (3)}$$

Where $Logit(P) = \ln[p/1-p]$, β is a vector of the log HRs for censoring among patients with similar baseline covariate history.

3.1.4.2 Weights based on baseline covariates and treatment at randomisation

The probability to crossover was modeled as a function of baseline covariates and treatment a patient received at randomisation. This is a more realistic approach as compared to 3.2.1 since now the probability to crossover does not only depend on baseline covariates but also on treatment to which patients were assigned to at randomisation. The fitted model is of the form;

$$\text{log}itP[C_{ij} = 0 \mid X_i, Trt_i] = \beta_0 + \beta X_i + \gamma Trt_i \qquad \text{Model (4)}$$

Where $Logit(P)$ and β are defined as above, Trt_i represents randomised treatment for patient i whereas γ is the log HR for censoring comparing randomised experimental treatment to control treatment among patients with same base line covariates.

3.1.4.3 Weights based on baseline covariates and all treatments taken by a patient

This approach allows the probability to crossover to not only depend on baseline covariates and treatment assigned at randomization but also takes into account the treatment the patient has taken up to the current time point j. This implies that if a patient has previously crossed from one treatment to another, they probably have less probability to switch treatment again at that time point. The conditional probabilities were fit using a logistic model below.

$$\text{log}itP[C_{ij} = 0 \mid X_i, A_{ij}] = \beta_0 + \beta X_i + \gamma A_{ij} \qquad \text{Model (5)}$$

Where $Logit(P)$ and β are defined as above whereas A_{ij} is treatment vector containing treatment taken by i^{th} patient at time points preceding and including j, γ is the log HR for

censoring taking into account the treatment a patient has taken up to time point j among patients with same base line covariates X .

3.1.4.4 Weights based on baseline covariates and treatment history

The final approach for calculating weights is one that is based on the probability of remaining uncensored until time point j conditioned on the measured baseline covariates and treatment history. Treatment history refers to the treatment the patient has taken up to time point $j-1$. These probabilities were obtained by fitting the following model;

$$\log itP[C_{ij}=0 \mid X_i, \bar{A}_{ij-1}] = \beta_0 + \beta X_i + \gamma \bar{A}_{ij-1} \qquad \text{Model (6)}$$

Where $Logit(P)$ and β are defined as before whereas $\bar{A}_{ij-1}$ is treatment history, γ is the log HR for censoring taking into account patients' treatment history among those with same base line covariates X (Cain & Cole, 2009)

For the third step of the IPCW analysis, for each set of the weights described above a weighted logistic regression model was fit for the risk of dying or disease progression. The model is of the form;

$$\log itP[D_{ij}=1 \mid X_i, Trt_i] = \beta_0 + \beta X_i + \gamma Trt_i \qquad \text{Model (7)}$$

Where D_{ij} =1 if the patient i experienced event at time point j and 0 if patient i did not experience event at time point j

This method has an advantage of being easily programmed in many standard statistical packages. However the use of weights induces within-subject correlation, which invalidates the standard error estimates output by a standard logistic program (they can be either too large or too small). This difficulty was overcome by fitting the above weighted logistic model using generalized estimating equations program (i.e. option "repeated" in SAS Proc Genmod)

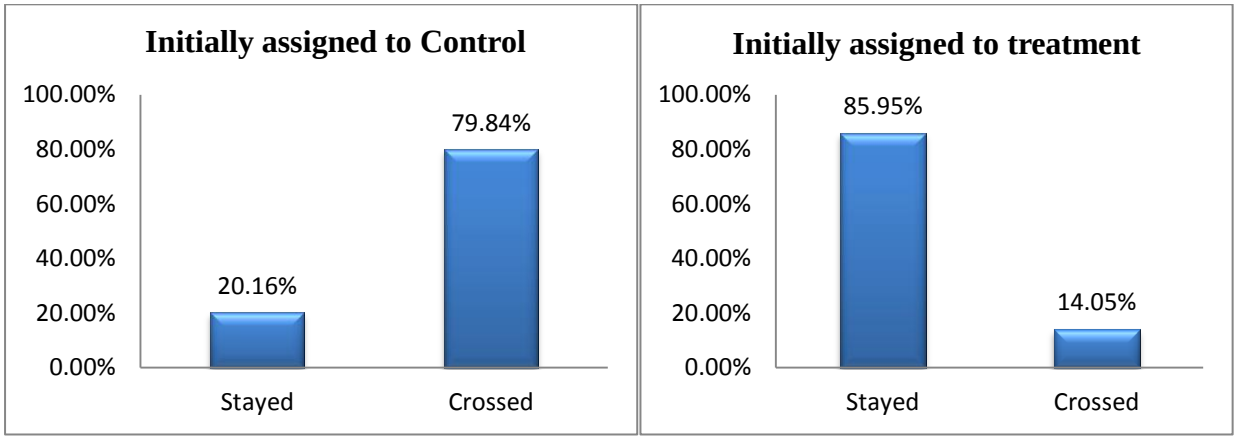
that outputs “robust” variance estimators that allow for correlated observations. (Hernan, Brumback, & Robins, 2000). A sample of the SAS program used is provided in Appendix B.

4 Results

4.1 Cross over statistics

Of the 243 patients enrolled in the control treatment group, 194 (79.84%) selectively crossed over to another treatment. Whereas only 14.05% of patients initially assigned to the experimental treatment crossed to another treatment. Switching of treatment confounds the interpretation of long-term follow-up data and raises the issue of how to best deal with compromised randomization in general. The proportions of crossovers for each treatment arm are shown in figure 1.

Figure 1: Crossover Statistics

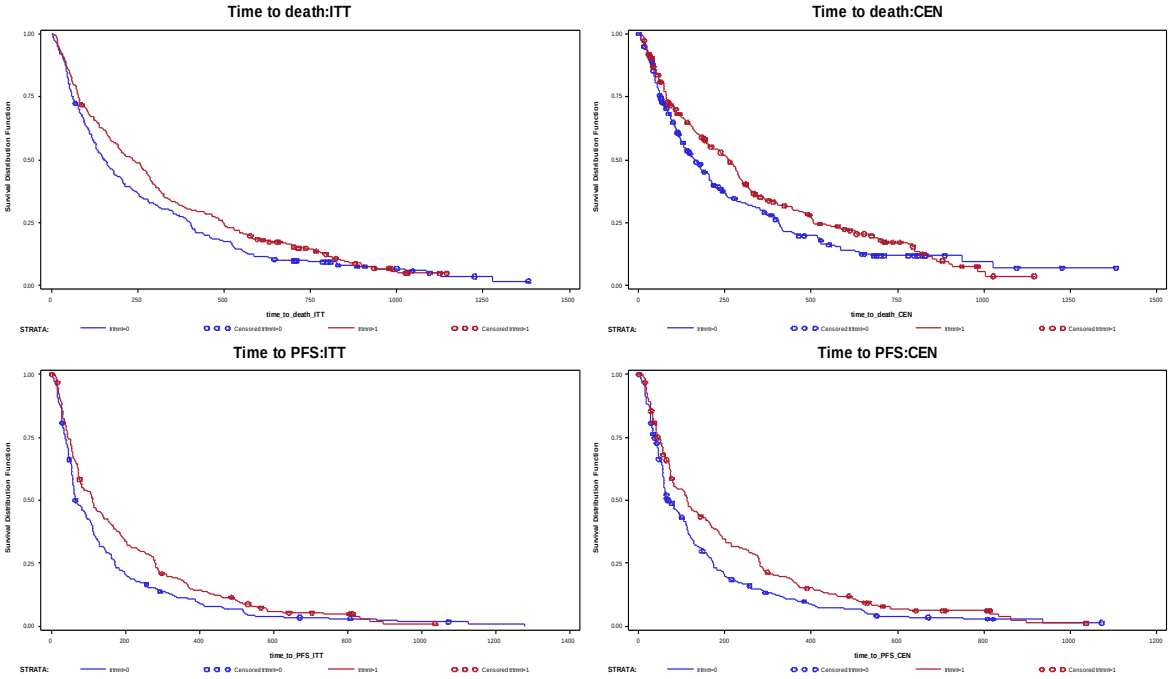


4.2 The Kaplan Meier estimate of the survival function

The Kaplan Meier was used to estimate the survival function and they give an insight into the shape of the survival function for each treatment arm. The Kaplan–Meier estimate of the survival function is a non-parametric method of estimating survival function from non- or right-censored data. This method is popular as it requires only very weak assumptions and yet utilises the information content of both fully observed and right-censored data. The Kaplan Meier estimate drops only at times when a failure has been observed.

Survival time distributions for intent to treat and censored at time of crossover data were estimated for each treatment arm as shown in figure 2. The curves seem to cross at the end; an indication of violation of proportional hazards assumption but this could be as results of patients switching treatments. It should also be noted that for the censored at time of crossover data, there are few event which remain at later time points after censoring the patients that switch treatment.

Figure 2: Kaplan Meier Curves



4.3 Statistical analysis

4.3.1 Overall survival

For the endpoint of overall survival, the results of the analysis are presented in figure 3 and detailed results are presented in appendix A. Table 4 provides a summary methods key describing the approach taken in each analysis. Baseline covariates described in section 2 were included for all the models fitted apart from the naïve model. Methods not adjusting for treatment crossover were considered in order to obtain a range of estimates of the treatment effect for comparison purposes.

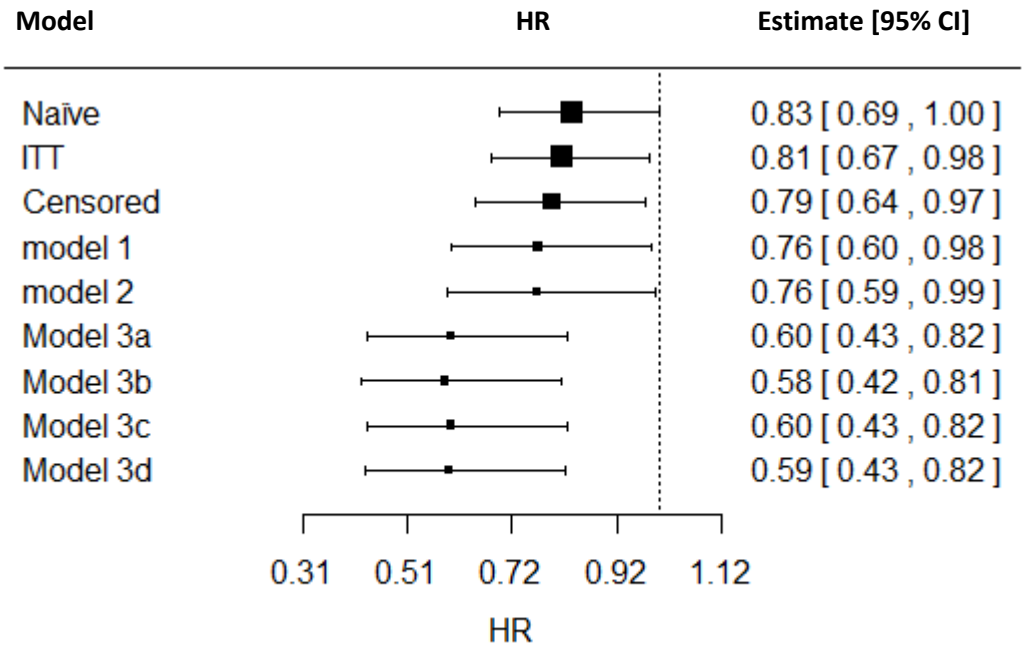
Table 4: Summary methods key

Method	***Description***
Naive	Cox proportional hazards model with treatment at baseline as the only covariate
ITT	Cox proportional hazards model based on Intent to treat
Censored	Cox proportional hazards model based on censored at crossover
Model1	Logistic model with outcome as dead or alive given baseline covariates and randomised treatment
Model2	Logistic model with outcome as dead or alive given baseline covariates and treatment history
Model3a	Logistic model with outcome as dead or alive given baseline covariates and randomised treatment with weights based on baseline covariates only
Model3b	Logistic model with outcome as dead or alive given baseline covariates and randomised treatment with weights based on baseline covariates and randomised treatment
Model3c	Logistic model with outcome as dead or alive given baseline covariates and randomised treatment with weights based on baseline covariates and treatment taken up to time point j
Model3d	Logistic model with outcome as dead or alive given baseline covariates and randomised treatment with weights based on baseline covariates and treatment history

The most naïve model that only includes treatment at randomisation produced a HR of 0.831 (95% CI 0.690, 1.001). In the ITT analysis, using a standard cox proportional hazards model with baseline covariates, the hazard ratio in favour of the experimental treatment was 0.813 (95% CI 0.673, 0.981). In ITT analysis, the efficacy attributed to the experimental treatment may have been greater than if all the patients had received experimental treatment as assigned. To account for the crossover, a censored analysis was undertaken. A cox regression model was fitted to censor at time of crossover data and a hazard ratio of 0.791 (95% CI 0.644, 0.973) was obtained.

Censoring crossover patients led to an increase in the estimate of the treatment effect. However due to informative censoring and large proportion of patients changing treatment, results obtained from these methods are prone to bias.

Figure 3: Results for Overall survival models



To adjust for confounding due to the informative censoring as a result of crossover, the parameters of the treatment effect were estimated by fitting a weighted cox proportional hazards model. Several time dependent weights were used varying according to the linear predictor in modelling the probability to crossover as discussed in section 3. The point estimates and 95% confidence intervals for each of the parameters are presented in figure 3. The estimates for the weighted models (3a-3d) range from 0.58 to 0.60 in favour of the experimental treatment. It can be noted that all the methods present a consistent message:- there is an overall survival advantage associated with experimental treatment. The naïve model gave a boundary result though this is because predicators other than treatment were not corrected for in the analysis. The results of IPCW method are indicative of a higher treatment effect compared to those obtained by methods not accounting for treatment crossover. However the exact size of the treatment effect is uncertain.

4.3.2 Crossover probabilities and weights for OS

The probabilities to crossover used to obtain weights were calculated by means of a logistic regression model as described in the previous section. Logistic models were fit for the outcome "crossover" which was a binary outcome indicating whether a patient stayed on the treatment they were assigned to or they changed to another study treatment. This model was fitted under different scenarios i.e. considering baseline covariates only, baseline covariates and treatment at baseline, baseline covariates and treatment taken up to time point t or baseline covariates and treatment history. Figure 4 shows the probabilities to crossover under the different scenarios for arbitrary selected patients. From the plots, the probability for patients to switch from one treatment to another increases overtime. This could be attributed to the fact that over time patients who tend to do badly are switched to the alternative treatment which shows to perform better or a non-study treatment all together for ethical reasons.

The box plots of the weights are presented in figure 5 for all the time points. The weights show a similar behaviour:-they decrease overtime. The distribution of the weights is skewed at time points 1 2 and 3 whereas their variance decreases with time. There were no outliers observed in the weights.

Figure 4: Probability to crossover for selected patients under Overall survival

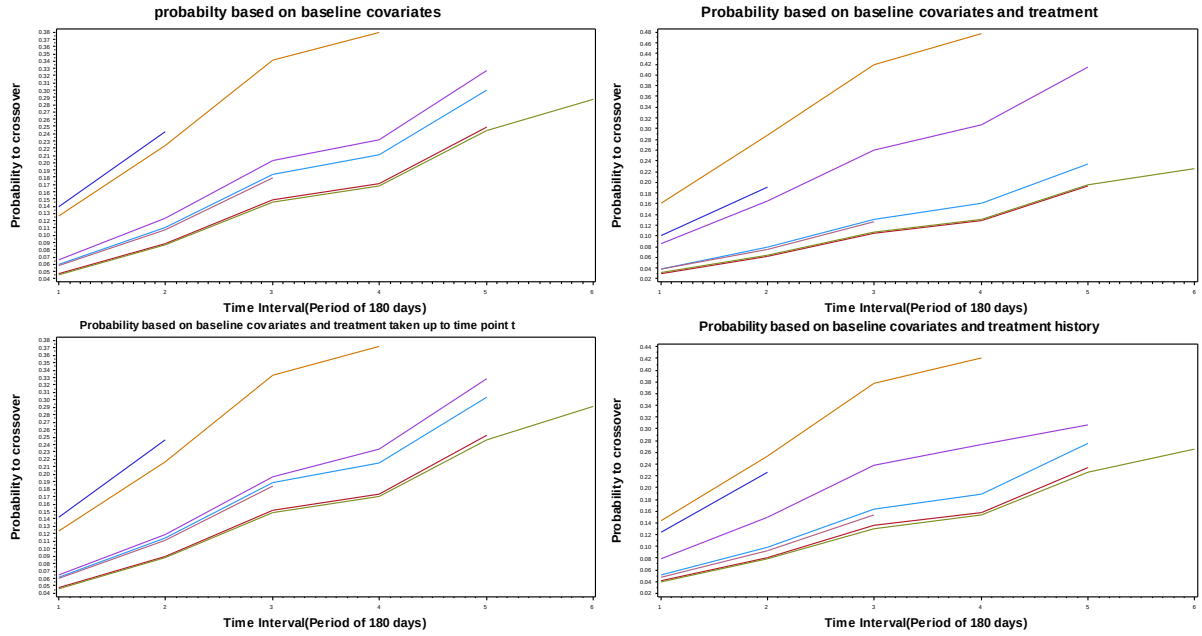
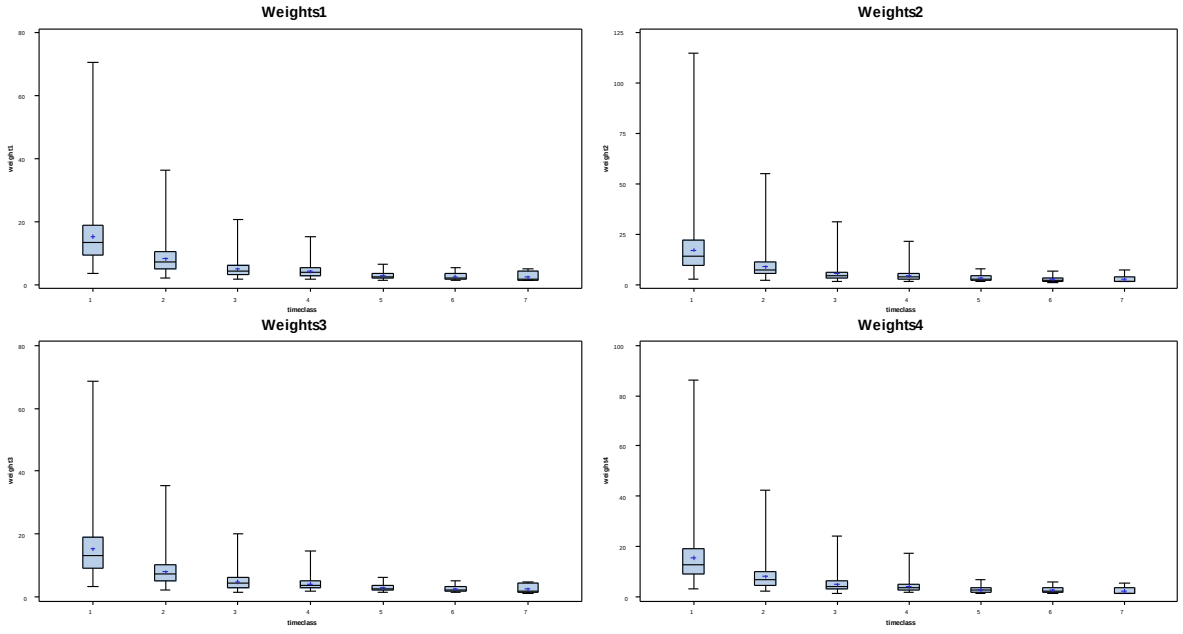


Figure 5: Boxplots for different weights par time point under OS

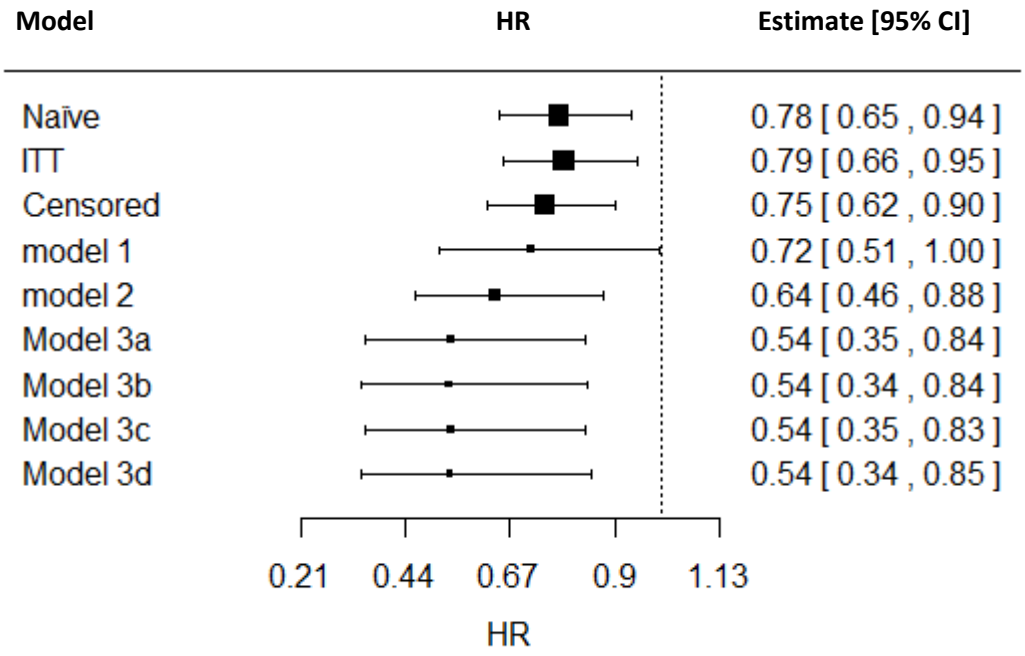


Weights1 correspond to weights calculated based on baseline covariates only; Weight2 corresponds to weights calculated based on baseline covariates and treatment; Weights3 corresponds to weights calculated based on baseline covariates and treatment taken up to time point j*; Weights4 corresponds to weights calculated based on baseline covariates and treatment taken up to time point* $j-1$.

4.3.3 Progression free survival

Similar models as described in section 3 were used on time to disease progression since there are patients who switched treatment before experiencing disease progression. Figure 6 presents the results of the various models fitted.

Figure 6: Results for Progression free survival models



Despite some of the patients in the control arm receiving the experimental treatment, intent-to-treat (ITT) analysis after follow-up still demonstrated a significantly longer PFS than the control treatment. The HR in favour of the experimental arm was, 0.79 (95% CI 0.66, 0.95) Results from the censored analysis show a slight improvement in the experimental treatment effect compared to ITT analysis.

Similar adjustments were done for PFS as for OS by fitting weighted cox regression model with weights obtained as discussed in section 3. The weighted models produced similar estimates of 0.54 with varying confidence limits in favour of the experimental treatment. The

experimental treatment effect observed from the weighted models is higher than that from naïve, ITT and censored analyses.

4.3.4 Crossover probabilities and weights for PFS

The probabilities to crossover used to obtain weights were calculated by means of a logistic regression model as described in the previous section. Figure 7 shows the probabilities to crossover under the different scenarios for arbitrary selected patients. From the plots, the probability for patients to switch from one treatment to another slightly increases overtime but not as high as those obtained from overall survival. This could be attributed to the fact that most patients experience disease progression before switching treatment.

The box plots of the weights are presented in figure 8 for all the time points. The weights show a similar behaviour:-they decrease overtime. The distribution of the weights is skewed at time points 1 2 and 3 whereas their variance decreases with time. There were no out liers observed in the weights. Patients who don't cross and have similar history as those that crossover get bigger weights i.e. these patients are "counted" more than once to make up for people like them that were censored due to crossover. Thus the reweighted population is no longer a biased sample.

Figure 7: Probability to crossover for selected patients under PFS

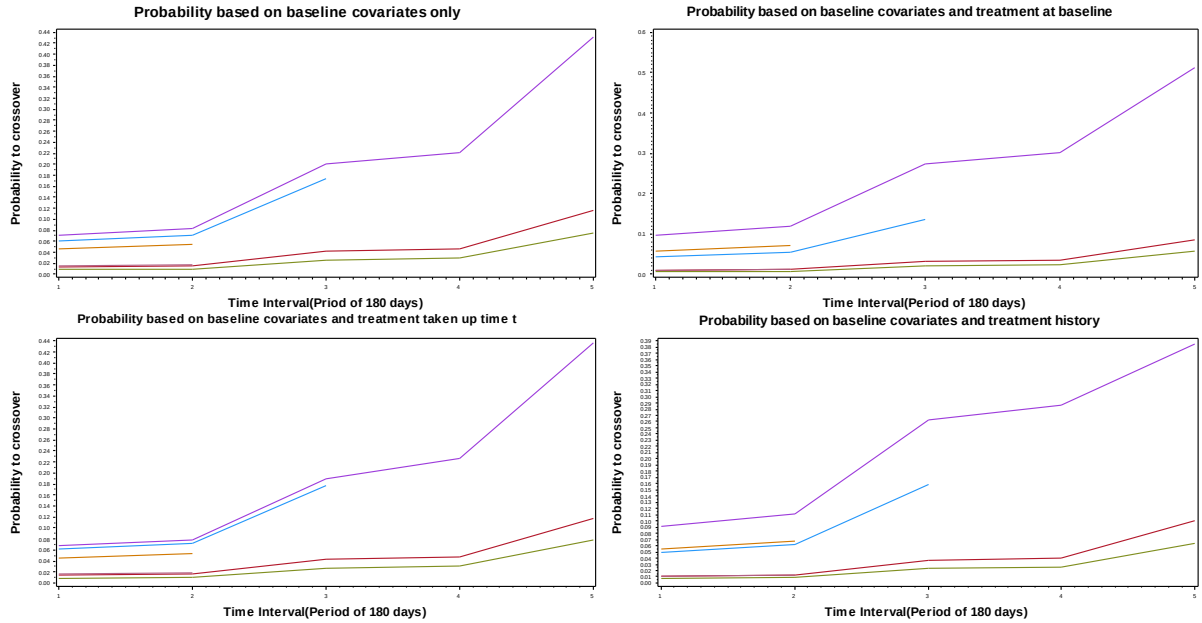
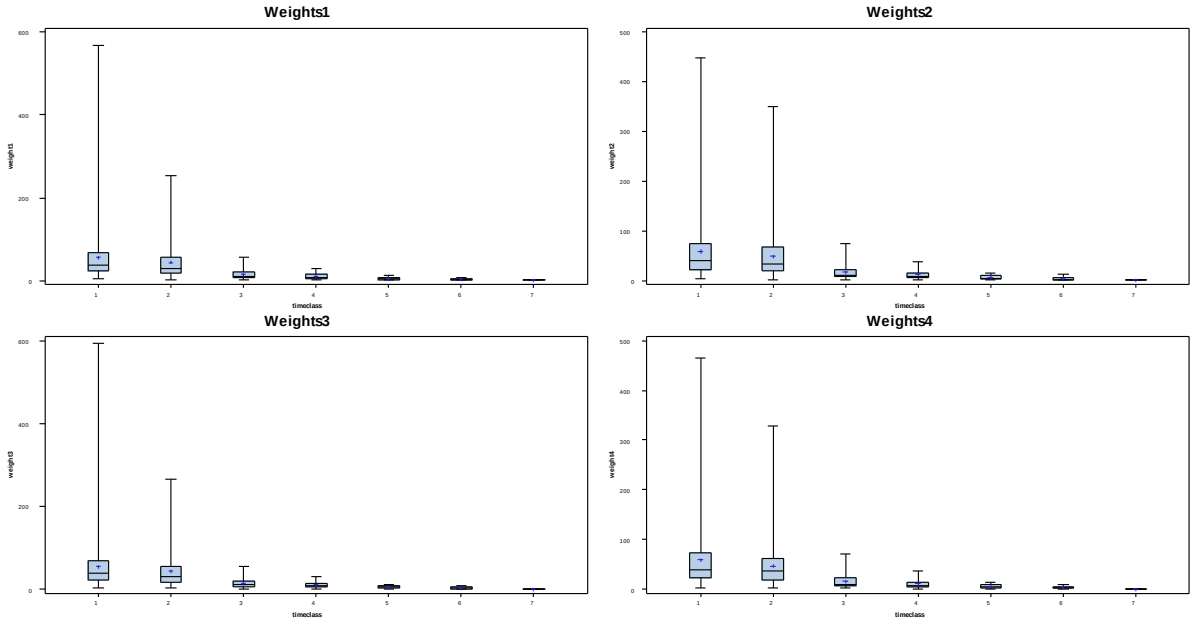


Figure 8: Boxplots for different weights per time point under PFS



Weights1 correspond to weights calculated based on baseline covariates only; Weight2 corresponds to weights calculated based on baseline covariates and treatment; Weights3 corresponds to weights calculated based on baseline covariates and treatment taken up to time point j*; Weights4 corresponds to weights calculated based on baseline covariates and treatment taken up to time point* $j-1$.

5 Discussion and conclusion

In the analysis, several methods were applied for estimating treatment effect in the presence of treatment crossover. For both overall survival and progression free survival data was transformed into a longitudinal sequence in order to fit weighted logistic models to estimate the treatment effect.

Time to progression outcome measures are not normally affected by treatment crossover in cases where crossover is only permitted after disease progression has occurred. This was not the situation in this study as some of the patients switched treatment before disease progression thus the survival estimates were adjusted for treatment crossover for this end point as well.

First, methods not accounting for crossover were applied i.e. ITT and censored analysis in order to obtain a range of estimates of the treatment effect for comparison purposes. The results from these methods are usually biased in the presence of treatment crossover (Latimer, et al.).

In ITT analysis, the survival data was analysed according to the arms to which patients were randomised. This implies that the estimate of the survival advantage associated with the experimental treatment will be biased. The bias in censored analysis arises from the fact that patient's survival times are censored at the time of crossover. This means that censoring is informative and methods of estimation not accounting for informative censoring can no longer be trusted to produce reliable results.

In order to adjust survival estimates in the presence of treatment crossover, the inverse probability of censoring weights (IPCW) method was applied to both overall survival and progression free survival data as discussed in the previous section. Several ways of estimating the weights were considered i.e. weights from probability of crossover given baseline

covariates only, given baseline covariates and treatment at baseline, given baseline covariates and treatment taken up to time point j, as well as baseline covariates and treatment history.

The weights were estimated to represent the inverse probability of informative censoring given factors affecting likelihood of crossover and/or survival. With these weights logistic regression models were fitted with event of death or disease progression as the dependent variable.

Because of the presence of confounding, the hazard ratios obtained from ITT and censored analyses were biased downwards. This indicated less overall survival advantage for the experimental treatment which was higher in the weighted analysis that provides under our assumption an unbiased estimate. The difference between un-weighted and weighted estimates could be attributed to the amount of confounding due to treatment crossover.

For all the scenarios, the weights showed a similar trend over time-the fact that the probability to cross over increases with time. Furthermore all methods presented a consistent message that there is an overall survival as well as progression free survival advantage associated with experimental treatment. However the size of the experimental treatment effect is uncertain. The results from IPCW method showed a slightly higher experimental treatment effect compared to that obtained from the ITT and censored analyses.

This result is expected because in this study 79.04% of control group patients switched treatment. When this happens, control group patients that crossover benefit from the experimental treatment and measures of overall survival in the control group will be higher that what would have been observed if treatment crossover had not occurred. This results in the overall survival advantage of the experimental treatment being underestimated.

IPCW analysis may be particularly useful for evaluating OS benefits that otherwise would be biased during ITT analysis. OS events are less frequent and often occur after disease progression; thus, evaluation of OS is affected to a greater degree by selective crossover than evaluation of progression free survival.

Although apparently useful, a limitation of the IPCW approach is that the variables capturing the important relationships between probability of censoring and probability of disease outcome (i.e., PFS or OS) must be known, or at least mostly known, and data must be available. If factors that predict outcome are absent, or if the determinants of crossover are unmeasured but nonetheless prognostic factors, then the adjustment will not produce valid estimates.

In conclusion, the application of inverse probability of censoring weights was described and applied to compare overall survival (OS) in patients 65 years and older who had newly diagnosed de novo or secondary AML and either poor or intermediate-risk cytogenetic who were randomly assigned to receive experimental treatment and control treatment. The results show significant longer survival for patients on experimental treatment. Further, the results show that correcting for bias due to crossover increases the treatment effect. Although this method is applied on a randomized trial, it can also be applicable to observational studies with time-varying treatments.

6 References

Howe, C. J., Cole, S. R., Chmiel, J. S., & Munoz, A. (2011). Limitation of Inverse Probability-of-Censoring Weights in Estimating Survival in the Presence of Strong Selection Bias. *American Journal of Epidemiology, 173*(5), 569-77.

Rimawi, M., & Hilsenbeck, S. G. (2012). Making Sense of Clinical Trial Data: Is Inverse Probability of Censoring Weighted Analysis the Answer to Crossover Bias? *Journal of Clinical Encology, 30*(4), 453-8.

Toh, S., Hernández-Díaz, S., Logan, R., Robins, J. M., & Hernán, M. A. (2010). Estimating absolute risks in the presence of nonadherence: An application to a follow-up study with baseline randomization. *21*(4), 528-539.

Cain, L. E., & Cole, S. R. (2009). Inverse probability-of-censoring weights for the correction of time-varying noncompliance in the effect of randomized highly active antiretroviral therapy on incident AIDS or death. *Statistics in Medicine*, 1725–1738.

Cole, S. R., & Cain, L. E. (2009). Inverse probability-of-censoring weights for the correction of time-varying noncompliance in the effect of randomized highly active antiretroviral therapy on incident AIDS or death. *Statistics in Medicine, 28*(12), 1725-38.

Hernan, M. A., Brumback, B., & Robins, J. M. (2000). Marginal Structural Models to Estimate the Causal Effect of Zidovudine on the Survival of HIV-positive Men. *Pubmed, 11*(5), 561-70.

Latimer, N., Wailoo, A., Abrams, K., Lambert, P., Morden, J., & Crowther, M. (n.d.). *Methods for adjusting survival estimates in the presence of treatment crossover-a simulation study and real world data analysis.*

Law, M. G., & Kaldor, J. M. (1996). Survival analysis of Randomized Clinical Trials adjusted for patients who switch treatments. *Statistics in Medicine, 15*(19), 2069-76.

Morden, J. P., Lambert, P. C., Latimer, N., Abrams, K. R., & Wailoo, A. J. (2011). Assessing methods for dealing with treatment switching in randomised controlled trials: a simulation study. *BMC Medical research Methodology, 11*(4). doi:10.1186/1471-2288-11-4

7 Appendix A (Additional results)

Appendix 1: ITT for OS

Parameter	*Estimate*	*SE*	*P-value*	*HR*	*95% HR Confidence Limits*
Trtmnt	-0.207	0.096	0.031	0.813	(0.673,0.981)
Age	0.023	0.009	0.007	1.024	(1.006,1.041)
ECOG(0-1)	-0.356	0.109	0.001	0.700	(0.566,0.867)
CGrisk(poor)	0.374	0.100	0.000	1.454	(1.196,1.767)
AMLtype(de novo)	0.063	0.101	0.533	1.065	(0.873,1.299)
Gender(male)	0.079	0.099	0.421	1.083	(0.892,1.314)
Race(white)	-0.006	0.143	0.966	0.994	(0.751,1.315)
BMBLAST(<=30%	-0.195	0.107	0.070	0.823	(0.667,1.016)

Appendix 2: Censored for OS

Parameter	*Estimate*	*SE*	*P-value*	*HR*	*95% HR Confidence Limits*
Trtmnt	-0.234	0.105	0.026	0.791	(0.644,0.973)
Age	0.027	0.009	0.005	1.027	(1.008,1.046)
ECOG(0-1)	-0.422	0.117	0.000	0.656	(0.521,0.825)
CGrisk(poor)	0.441	0.109	<.0001	1.554	(1.256,1.924)
Gender(male)	0.133	0.109	0.221	1.142	(0.923,1.414)
Race(white)	-0.145	0.156	0.354	0.865	(0.637,1.175)
BMBLAST(<=30%)	-0.120	0.116	0.299	0.887	(0.707,1.113)
AMLtype(de novo)	0.158	0.111	0.155	1.172	(0.942,1.458)

Appendix 3: Model1 for OS

Parameter	*Estimate*	*Standard Error*	*95% Confidence Limits*	*P-value*
Intercept	-2.647	1.076	(-4.7563,-0.5376)	0.014
Trtmnt	-0.268	0.128	(-0.5189,-0.0169)	0.036
Age	0.028	0.013	(0.0019,0.0541)	0.036
ECOG(0-1)	0.380	0.159	(0.0693,0.6911)	0.017
CGrisk(poor)	-0.433	0.149	(-0.7243,-0.1418)	0.004
AMLtype(de novo)	-0.058	0.134	(-0.3198,0.2038)	0.664
Gender(male)	-0.097	0.128	(-0.3475,0.1531)	0.447
Race(white)	0.000	0.186	(-0.365,0.3649)	1.000
BMBLAST(<=30%)	0.197	0.141	(-0.0793,0.4732)	0.162
Timeclass1	0.800	0.655	(-0.4848,2.0839)	0.222
Timeclass2	0.501	0.665	(-0.8017,1.8034)	0.451
Timeclass3	0.552	0.672	(-0.765,1.8694)	0.411
Timeclass4	0.162	0.697	(-1.2037,1.5269)	0.817
Timeclass5	0.210	0.741	(-1.2423,1.6613)	0.777
Timeclass6	-0.138	0.836	(-1.7756,1.5004)	0.869

Appendix 4: Model2 for OS

Parameter	*Estimate*	*SE*	*95% Confidence Limits*	*P-value*
Intercept	-2.480	1.114	(-4.6629,-0.2965)	0.026
A	-0.270	0.134	(-0.5325,-0.0083)	0.043
Age	0.027	0.013	(0.0012,0.0531)	0.040
ECOG(0-1)	0.382	0.161	(0.0667,0.6972)	0.018
CGrisk(poor)	-0.450	0.148	(-0.7392,-0.1601)	0.002
AMLtype(de novo)	-0.044	0.134	(-0.3073,0.219)	0.742
Gender(male)	-0.090	0.128	(-0.3411,0.1607)	0.481
Race(white)	-0.006	0.190	(-0.3791,0.3674)	0.976
BMBLAST(<=30%)	0.181	0.140	(-0.0933,0.4554)	0.196
Timeclass1	0.733	0.692	(-0.6225,2.0883)	0.289
Timeclass2	0.434	0.698	(-0.9336,1.8009)	0.534
Timeclass3	0.487	0.706	(-0.8967,1.8697)	0.491
Timeclass4	0.092	0.725	(-1.3297,1.5135)	0.899
Timeclass5	0.158	0.762	(-1.3357,1.651)	0.836
Timeclass6	-0.186	0.844	(-1.8406,1.4678)	0.825

Appendix 5: model 3a for OS

Parameter	*Estimate*	*SE*	*95% Confidence Limits*	*P-value*
Intercept	-3.634	1.365	(-6.3091,-0.9597)	0.008
Trtmnt	-0.516	0.163	(-0.8346,-0.1965)	0.002
ECOG(0-1)	0.656	0.196	(0.273,1.0399)	0.001
Age	0.029	0.015	(-0.0008,0.0593)	0.057
BMBLAST(<=30%)	0.262	0.177	(-0.0846,0.6087)	0.139
Race(white)	0.205	0.235	(-0.2565,0.6658)	0.384
Gender(male)	-0.043	0.164	(-0.3637,0.2781)	0.794
CGrisk(poor)	-0.606	0.177	(-0.9536,-0.2586)	0.001
AMLtype(de novo)	-0.076	0.170	(-0.4096,0.2584)	0.657
Timeclass1	1.385	0.754	(-0.0923,2.8628)	0.066
Timeclass2	1.117	0.758	(-0.369,2.6035)	0.141
Timeclass3	1.224	0.762	(-0.269,2.7177)	0.108
Timeclass4	0.736	0.776	(-0.7849,2.2572)	0.343
Timeclass5	0.961	0.793	(-0.5931,2.5152)	0.226
Timeclass6	0.501	0.873	(-1.2109,2.2129)	0.566

Appendix 6: Model 3b for OS

Parameter	*Estimate*	*SE*	*95% Confidence Limits*	*P-value*
Intercept	-2.918	1.428	(-5.7172,-0.1184)	0.041
Trtmnt	-0.537	0.166	(-0.8618,-0.2121)	0.001
Age	0.021	0.017	(-0.0112,0.0535)	0.200
ECOG(0-1)	0.588	0.213	(0.1706,1.0051)	0.006
CGrisk(poor)	-0.624	0.184	(-0.9853,-0.2623)	0.001
AMLtype(de novo)	-0.078	0.178	(-0.426,0.2697)	0.660
Gender(male)	-0.065	0.175	(-0.4068,0.2776)	0.712
Race(white)	0.156	0.236	(-0.3074,0.6188)	0.510
BMBLAST(<=30%)	0.233	0.188	(-0.1349,0.6015)	0.214
Timeclass1	1.528	0.737	(0.0832,2.9726)	0.038
Timeclass2	1.337	0.743	(-0.1201,2.7938)	0.072
Timeclass3	1.312	0.751	(-0.1606,2.7835)	0.081
Timeclass4	0.813	0.762	(-0.6801,2.3067)	0.286
Timeclass5	1.270	0.785	(-0.269,2.8085)	0.106
Timeclass6	0.747	0.876	(-0.9708,2.4644)	0.394

Appendix 7: Model 3c for OS

Parameter	*Estimate*	*SE*	*95% Confidence Limits*	*P-value*
Intercept	-3.716	1.368	(-6.3963,-1.035)	0.007
Trtmnt	-0.516	0.163	(-0.8344,-0.1974)	0.002
Age	0.030	0.015	(-0.0001,0.0599)	0.051
ECOG(0-1)	0.661	0.196	(0.2781,1.0446)	0.001
CGrisk(poor)	-0.604	0.178	(-0.9519,-0.2556)	0.001
AMLtype(de novo)	-0.075	0.171	(-0.4098,0.2597)	0.660
Gender(male)	-0.038	0.164	(-0.3587,0.283)	0.817
Race(white)	0.206	0.237	(-0.2583,0.6701)	0.385
BMBLAST(<=30%)	0.258	0.177	(-0.0883,0.6047)	0.144
Timeclass1	1.404	0.756	(-0.0786,2.8864)	0.063
Timeclass2	1.128	0.761	(-0.3629,2.6189)	0.138
Timeclass3	1.241	0.764	(-0.256,2.7388)	0.104
Timeclass4	0.752	0.778	(-0.773,2.2767)	0.334
Timeclass5	0.966	0.794	(-0.5895,2.5221)	0.224
Timeclass6	0.522	0.873	(-1.1889,2.2331)	0.550

Appendix 8: model 3d for OS

Parameter	*Estimate*	*SE*	*95% Confidence Limits*	*P-value*
Intercept	-3.239	1.370	(-5.9239,-0.5545)	0.018
Trtmnt	-0.521	0.165	(-0.8439,-0.1988)	0.002
ECOG(0-1)	0.630	0.200	(0.2379,1.022)	0.002
Age	0.026	0.016	(-0.0049,0.0566)	0.099
BMBLAST(<=30%)	0.266	0.180	(-0.0864,0.6189)	0.139
Race(white)	0.186	0.231	(-0.2672,0.6393)	0.421
Gender(male)	-0.058	0.167	(-0.3845,0.2683)	0.727
CGrisk(poor)	-0.606	0.178	(-0.9548,-0.2571)	0.001
AMLtype(de novo)	-0.070	0.171	(-0.405,0.2658)	0.684
Timeclass1	1.299	0.745	(-0.1616,2.7601)	0.081
Timeclass2	1.068	0.751	(-0.4034,2.5384)	0.155
Timeclass3	1.131	0.756	(-0.3508,2.6128)	0.135
Timeclass4	0.660	0.771	(-0.85,2.1703)	0.392
Timeclass5	0.940	0.789	(-0.6065,2.4865)	0.234
Timeclass6	0.399	0.879	(-1.3238,2.122)	0.650

Appendix 9: Results for probability to crossover corresponding to Model 3d for OS

Parameter	*Estimate*	*SE*	*P-value*
Intercept	4.339	1.509	0.004
B	-0.323	0.181	0.074
ECOG(0-1)	-0.061	0.221	0.782
Age	-0.082	0.018	<.0001
BMBLAST(<=30%)	0.277	0.206	0.179
Race(white)	0.711	0.238	0.003
Gender(male)	0.148	0.186	0.428
CGrisk(poor)	-0.392	0.191	0.040
AMLtype(de novo)	-0.045	0.192	0.815
Timeclass1	-1.407	0.198	<.0001
Timeclass2	-0.690	0.205	0.001
Timeclass3	-0.112	0.216	0.603
Timeclass4	0.070	0.250	0.781
Timeclass5	0.556	0.296	0.061
Timeclass6	0.768	0.398	0.053

Appendix 10: ITT for PFS

Parameter	*Estimate*	*SE*	*P-value*	*HR*	*95% HR CI*
Trtmnt	-0.237	0.094	0.012	0.789	(0.656,0.949)
ECOG(0-1)	-0.269	0.107	0.012	0.764	(0.62,0.942)
Age	0.010	0.009	0.258	1.010	(0.993,1.027)
Gender(male)	0.253	0.097	0.009	1.288	(1.064,1.559)
BMBLAST(<=30%)	-0.094	0.105	0.371	0.911	(0.742,1.118)
Race(white)	-0.065	0.144	0.652	0.937	(0.707,1.242)
CGrisk(poor)	0.304	0.099	0.002	1.355	(1.117,1.644)
AMLtype(de novo)	0.076	0.100	0.448	1.079	(0.887,1.313)

Appendix 11: Censored for PFS

Parameter	*Estimate*	*SE*	*P-value*	*HR*	*95% HR CI*
Trtmnt	-0.291	0.096	0.003	0.747	(0.619,0.903)
ECOG(0-1)	-0.273	0.108	0.011	0.761	(0.616,0.94)
Age	0.006	0.009	0.498	1.006	(0.989,1.023)
Gender(male)	0.251	0.100	0.012	1.285	(1.057,1.563)
BMBLAST(<=30%)	-0.049	0.106	0.642	0.952	(0.773,1.172)
Race(white)	-0.208	0.145	0.151	0.812	(0.611,1.079)
CGrisk(poor)	0.367	0.101	0.000	1.443	(1.183,1.76)
AMLtype(de novo)	0.070	0.102	0.495	1.072	(0.878,1.309)

Appendix 12: Model1 for PFS

Parameter	*Estimate*	*SE*	*95% CI*	*P-value*
Intercept	-0.070	1.351	(-2.7164,2.5774)	0.959
Trtmnt	-0.333	0.170	(-0.6657,-0.001)	0.049
ECOG(0-1)	0.301	0.199	(-0.0894,0.6923)	0.131
Age	0.022	0.017	(-0.0113,0.0552)	0.196
BMBLAST(<=30%)	0.165	0.176	(-0.18,0.5099)	0.349
Race(white)	0.057	0.276	(-0.4828,0.5975)	0.835
Gender(male)	-0.440	0.155	(-0.743,-0.1375)	0.004
CGrisk(poor)	-0.369	0.198	(-0.7565,0.0183)	0.062
AMLtype(de novo)	-0.207	0.158	(-0.5156,0.1016)	0.189
Timeclass1	0.288	0.691	(-1.066,1.6414)	0.677
Timeclass2	-0.278	0.720	(-1.6887,1.1332)	0.700
Timeclass3	-0.092	0.758	(-1.5775,1.3928)	0.903
Timeclass4	-0.656	0.800	(-2.224,0.9113)	0.412
Timeclass5	0.412	0.996	(-1.5413,2.3644)	0.680
Timeclass6	-0.140	1.314	(-2.7152,2.4343)	0.915

Appendix 13:model2 for PFS

Parameter	*Estimate*	*SE*	*95% Confidence Limits*	*P-value*
Intercept	0.292	1.366	(-2.3845,2.9682)	0.831
A	-0.451	0.163	(-0.7712,-0.1316)	0.006
ECOG(0-1)	0.306	0.200	(-0.0869,0.6984)	0.127
Age	0.022	0.017	(-0.0117,0.0549)	0.204
BMBLAST(<=30%)	0.166	0.175	(-0.1777,0.5097)	0.344
Race(white)	0.119	0.257	(-0.3846,0.6221)	0.644
Gender(male)	-0.430	0.156	(-0.7351,-0.1239)	0.006
CGrisk(poor)	-0.394	0.198	(-0.7807,-0.0066)	0.046
AMLtype(de novo)	-0.198	0.159	(-0.5095,0.1142)	0.214
Timeclass1	-0.044	0.691	(-1.3973,1.3099)	0.950
Timeclass2	-0.595	0.714	(-1.9934,0.8043)	0.405
Timeclass3	-0.411	0.753	(-1.8869,1.0656)	0.586
Timeclass4	-0.983	0.785	(-2.5218,0.5553)	0.210
Timeclass5	0.144	0.947	(-1.7108,1.9995)	0.879
Timeclass6	-0.344	1.237	(-2.7681,2.08)	0.781

Appendix 14: model3a for PFS

Parameter	*Estimate*	*SE*	*95% Confidence Limits*	*P-value*
Intercept	0.154	1.786	(-3.3462,3.6548)	0.931
Trtmnt	-0.612	0.222	(-1.0468,-0.1779)	0.006
ECOG(0-1)	0.403	0.236	(-0.0604,0.8659)	0.088
Age	0.026	0.019	(-0.0118,0.0637)	0.178
BMBLAST(<=30%)	0.025	0.240	(-0.4452,0.4944)	0.918
Race(white)	0.288	0.365	(-0.4263,1.003)	0.429
Gender(male)	-0.438	0.218	(-0.8641,-0.0109)	0.044
CGrisk(poor)	-0.712	0.245	(-1.193,-0.2316)	0.004
AMLtype(de novo)	-0.103	0.221	(-0.5364,0.3296)	0.640
Timeclass1	0.174	0.958	(-1.7025,2.051)	0.856
Timeclass2	-0.184	1.012	(-2.1667,1.7987)	0.856
Timeclass3	0.151	1.044	(-1.8962,2.1973)	0.885
Timeclass4	-1.365	1.057	(-3.4364,0.7057)	0.196
Timeclass5	0.435	1.189	(-1.8943,2.7649)	0.714
Timeclass6	1.398	1.602	(-1.7431,4.5381)	0.383

Appendix 15: model 3b for PFS

Parameter	*Estimate*	*SE*	*95% Confidence Limits*	*P-value*
Intercept	0.538	1.800	(-2.9892,4.0656)	0.765
Trtmnt	-0.622	0.227	(-1.0678,-0.1762)	0.006
ECOG(0-1)	0.350	0.244	(-0.1272,0.8277)	0.151
Age	0.021	0.020	(-0.0184,0.0601)	0.298
BMBLAST(<=30%)	0.075	0.242	(-0.3996,0.5504)	0.756
Race(white)	0.288	0.361	(-0.419,0.9949)	0.425
Gender(male)	-0.486	0.228	(-0.9332,-0.038)	0.034
CGrisk(poor)	-0.669	0.259	(-1.1751,-0.1619)	0.010
AMLtype(de novo)	-0.123	0.231	(-0.5754,0.3299)	0.595
Timeclass1	0.155	0.958	(-1.7219,2.0317)	0.872
Timeclass2	-0.122	1.011	(-2.1041,1.8594)	0.904
Timeclass3	0.128	1.051	(-1.9321,2.188)	0.903
Timeclass4	-1.270	1.047	(-3.323,0.7824)	0.225
Timeclass5	0.772	1.215	(-1.609,3.1521)	0.525
Timeclass6	1.757	1.705	(-1.5841,5.0975)	0.303

Appendix 16: model3c for PFS

Parameter	*Estimate*	*SE*	*95% Confidence Limits*	*P-value*
Intercept	0.080	1.795	(-3.4384,3.5985)	0.964
Trtmnt	-0.613	0.220	(-1.0438,-0.1814)	0.005
ECOG(0-1)	0.409	0.237	(-0.0544,0.8729)	0.084
Age	0.027	0.019	(-0.0109,0.0646)	0.163
BMBLAST(<=30%)	0.020	0.241	(-0.4516,0.4915)	0.934
Race(white)	0.292	0.370	(-0.4335,1.0167)	0.431
Gender(male)	-0.430	0.217	(-0.8557,-0.0048)	0.048
CGrisk(poor)	-0.719	0.245	(-1.199,-0.2384)	0.003
AMLtype(de novo)	-0.100	0.221	(-0.5337,0.3336)	0.651
Timeclass1	0.176	0.959	(-1.704,2.0556)	0.855
Timeclass2	-0.195	1.014	(-2.1819,1.7929)	0.848
Timeclass3	0.144	1.046	(-1.9055,2.1936)	0.890
Timeclass4	-1.368	1.062	(-3.4497,0.7138)	0.198
Timeclass5	0.401	1.187	(-1.9247,2.7269)	0.735
Timeclass6	1.400	1.586	(-1.7092,4.5082)	0.378

Appendix 17: model3d for PFS

Parameter	*Estimate*	*SE*	*95% CI*	*P-value*
Intercept	0.469	1.799	(-3.0563,3.9951)	0.794
Trtmnt	-0.619	0.231	(-1.071,-0.1666)	0.007
ECOG(0-1)	0.370	0.244	(-0.108,0.847)	0.129
Age	0.022	0.020	(-0.017,0.0606)	0.270
BMBLAST(<=30%)	0.060	0.242	(-0.4144,0.5345)	0.804
Race(white)	0.275	0.357	(-0.4251,0.9754)	0.441
Gender(male)	-0.474	0.227	(-0.919,-0.0299)	0.037
CGrisk(poor)	-0.683	0.254	(-1.1802,-0.1849)	0.007
AMLtype(de novo)	-0.109	0.228	(-0.5551,0.3372)	0.632
Timeclass1	0.157	0.957	(-1.7188,2.0331)	0.870
Timeclass2	-0.123	1.012	(-2.106,1.8591)	0.903
Timeclass3	0.160	1.049	(-1.8953,2.2152)	0.879
Timeclass4	-1.385	1.047	(-3.4361,0.6667)	0.186
Timeclass5	0.616	1.204	(-1.7442,2.9759)	0.609
Timeclass6	1.394	1.701	(-1.9388,4.7272)	0.412

8 Appendix B (SAS Code)

(1) Cox Proportional hazards regression model (code corresponds to model ITT for OS)

```
proc phreg data=summerp3 plots(cl)=survival;
class  ecog gender trtmnt(param=ref ref="0") bmblast race cgrisk amltype;
model time_to_death_ITT*status_death_ITT(0)=trtmnt age ecog cgrisk amltype
gender race bmblast /rl; run;
```

(2) Un-weighted model for probability of event given baseline covariates treatment history (Code corresponds to model2 for OS)

```
proc genmod data=time_to_death_ITT_final descending;
class id timeclass;
model history=A age ecog cgrisk amltype gender race bmblast timeclass/
link=logit dist=bin;
repeated subject=id/type=ind;
run;
```

(3) Code for model 3d for OS

```
data time_to_death_ITT_final;
set time_to_death_ITT_final;
B=lag (A);
by id;run;
```

(i) Obtaining probabilities to crossover

```
proc logistic data=time_to_death_ITT_final;
class id timeclass;
model c=B ecog age bmblast race gender cgrisk amltype timeclass;
output out=weight4 p=pred4;
run;
```

(ii) Calculating the weights from the probabilities to crossover

```
data weight4;
set weight4;
weight4=1/pred4;
run;
```

(iii) Fitting weighted model for probability of event given baseline covariates and treatment history with weights4 (Code corresponds to model 3d for OS)

```
proc genmod data=weight4 descending;
class id timeclass;
model D=trtmnt ecog age bmblast race gender cgrisk amltype timeclass/
link=logit dist=bin;
scwgt weight4;
repeated subject=id/type=ind;
run;
```

Auteursrechtelijke overeenkomst

Ik/wij verlenen het wereldwijde auteursrecht voor de ingediende eindverhandeling:
Correcting for cross-over bias in randomized controlled clinical trials

Richting: **Master of Statistics-Biostatistics**
Jaar: **2012**

in alle mogelijke mediaformaten, - bestaande en in de toekomst te ontwikkelen - , aan de Universiteit Hasselt.

Niet tegenstaand deze toekenning van het auteursrecht aan de Universiteit Hasselt behoud ik als auteur het recht om de eindverhandeling, - in zijn geheel of gedeeltelijk -, vrij te reproduceren, (her)publiceren of distribueren zonder de toelating te moeten verkrijgen van de Universiteit Hasselt.

Ik bevestig dat de eindverhandeling mijn origineel werk is, en dat ik het recht heb om de rechten te verlenen die in deze overeenkomst worden beschreven. Ik verklaar tevens dat de eindverhandeling, naar mijn weten, het auteursrecht van anderen niet overtreedt.

Ik verklaar tevens dat ik voor het materiaal in de eindverhandeling dat beschermd wordt door het auteursrecht, de nodige toelatingen heb verkregen zodat ik deze ook aan de Universiteit Hasselt kan overdragen en dat dit duidelijk in de tekst en inhoud van de eindverhandeling werd genotificeerd.

Universiteit Hasselt zal mij als auteur(s) van de eindverhandeling identificeren en zal geen wijzigingen aanbrengen aan de eindverhandeling, uitgezonderd deze toegelaten door deze overeenkomst.

Voor akkoord,

Natukunda, Agnes

Datum: **14/09/2012**