2011
2012

FACULTY OF SCIENCES

Master of Statistics: Biostatistics

Masterproef

Adjusting for crossover bias in an observational study for patients with multiple myeloma

Promotor :
Prof. dr. Ziv SHKEDY

Promotor :
Dr. SUZY VAN SANDEN

Aneth Vedastus Kalinjuma

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

Maastricht University

universiteit hasselt
KNOWLEDGE IN ACTION

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

2011
2012

FACULTY OF SCIENCES

Master of Statistics: Biostatistics

Masterproef

Adjusting for crossover bias in an observational study for patients with multiple myeloma

Promotor :
Prof. dr. Ziv SHKEDY

Promotor :
Dr. SUZY VAN SANDEN

Aneth Vedastus Kalinjuma

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

Maastricht University

universiteit hasselt
KNOWLEDGE IN ACTION

Acknowledgement

This project was supported with Janssen Pharmaceutical Company Belgium in collaboration with Hasselt University, statistics department. It's a project to fulfill the requirements of Master of Statistics (Biostatistics ICP). I real appreciate for this opportunity that Janssen Company has given me to work on this important data set, thank you very much. I would like to give thanks to my all supervisors Prof. Shkedy, Ziv, Van Sanden, Suzy and Diels, Joris for their great support during the entire period of working with them. Many thanks go to my internal supervisor Prof Shkedy, Ziv who was available all the time when I needed his support to accomplish the project.

I also send my gratitude to VLIR-UOS scholarship for awarding me this important opportunity for developing my carrier in statistics. This real improved my views in terms of my academic perspective and statistics as a whole. I'm now confident with high quality education that I have received at Hasselt University, and my next step is to implement it in the field. Thanks to VLIR-UO coordinator of Hasselt University Mrs Martine Machiels for all the support she had given me during my studies. Further many thanks go to all Professors of Statistics department for their effort to make sure I got the high quality Master of Statistics in Biostatistics.

I give Thanks to our Almighty God for extra ordinary strength he has been giving me during the course. It was real tough but through him I survived. This project is dedicated to my parents Mr. Vedastus Kalinjuma and Mrs. Aurelia Kalinjuma. For the memory of my further, I real remember how much it costs you to take me to school. Thank you very much; I will always remember all the support you provided to me to become who I'm now. To my lovely Mom, thank you for being strong all the time and encouragement you always give me all the time, to me you means a lot.

To Mr. Wilbard Kalinjuma, I real thank him for taking care of me from the time my father passed away, it was difficult for him but he was strong enough to continue being our further. Lastly but not least this project is dedicated to my sisters and brothers Edina Kalinjuma, Alena Kalinjuma, James Kalinjuma and Dickson Willum. Thank you for your support and being there for me.

Table of Content

List of Abbreviations

IPTW	Inverse Probability of Treatment Weights
Cox PH	Cox Proportional Hazard
OS	Overall Survival
PFS	Progression Free Survival
IPCW	Inverse Probability Censoring Weight
MSM	Marginal Structural Model
NSM	Nested Structural Model

Abstract

Estimation of treatment effect in observational studies is not straight forward due to lack of randomization. As a result blind comparison of treatment effects will lead to biased estimates. Another challenge that exists in observational studies is treatment crossover. In this project this was a major complication because some of patients were crossing between treatments of interest and after sometime they also crossed to other treatments too. Therefore analysis that does not take into account this issue will also result into biased estimates.

The objective of this study was to estimate treatment effect of Combination Therapy in patients with multiple myeloma in three scenarios: (1) Ignore all sources of bias and estimate treatment effect (2) Control selection bias due to the fact that it is an observational study (3) Control of treatment crossover bias in estimation of treatment effect.

There exist different methods for estimating treatment effect for observational studies. For this project ordinary models, propensity score methods, Censoring and weighted analysis were applied. Propensity score is a balancing score that is usually used in observational studies for balancing distribution of baseline and/or prognostic factors between treatment groups. After estimating these score different models can be fitted. In this project Cox proportional hazard model with stratified propensity score and weighted Cox PH model with (IPTW based on propensity scores) were applied. The censoring method is one of the traditional ways of dealing with treatment crossover. This method censors subjects/patients at the time they crossed treatment. There after treatment estimation was done in a normal way. For this analysis Cox PH model was fitted in order to estimate treatment effect. Finally the analysis based on weights was also fitted as a most plausible way of estimating treatment effect in presence of both selection bias and treatment crossover bias. The key idea for this weighted analysis is that; for those who are censored due to crossover will receive lower weight than those who stayed. By doing so it creates a population that would have been observed when there was no crossover. After estimating weights, weighted Cox PH model was fitted to estimate treatment effect. For this project three ways of estimating weights were applied to see their effect on the parameter of interest.

Results showed that Ordinary and propensity score methods were supporting standard treatment (Mono Therapy), whereas weighted models showed treatment effect for Combination Therapy. Moreover as weights were modified, the more effect was observed for combination therapy. This treatment effect was strong for OS as compared to PFS. To summarise, for this projects weighted analysis was preferred for controlling both selection bias and treatment crossover bias.

1. Introduction

Researchers in public health normally are interested in comparing treatments so that the treatment that performs better can be adopted. However comparing treatments in an observational study is not direct due to lack of randomization. As a result comparing treatments in this nature of data may simply reflect the underlying difference between treatment groups and not treatment effect (Curtis, Hammill, Eisenstein, Kramer, & Anstrom, 2007).

The main difference between experimental studies (randomized clinical trial) and observational studies is that; in experimental studies the assignment of the treatments to subjects in the study is controlled by the experimenter. He/she ensures that subjects receiving different treatments are comparable. Whereas an observational study investigator does not have control on the assignments of the treatments and therefore he/she cannot ensure that similar subjects receive different treatments (Rosenbaum, 2002).

Generally literatures shows observational studies give an important knowledge of disease and its cause in public health. Observational studies are preferred over experimental studies for different reasons namely (1) Ethics issues: in some studies it is unethical to conduct experimental study for example lung cancer studies. In this type of studies experimenter is not allowed to expose subjects to smoke so that he/she can be able to study causal effect between smoking and lung cancer (Rosenbaum, 2002). (2) Refusal: sometimes people may refuse to be assigned in control group for different reasons that experimenter cannot control as a result the experimental study cannot be conducted in such situation (Rosenbaum, 2002). (3) Rare diseases: conducting experimental studies for rare disease is not practical due to few cases of such disease. As result one opts for observational study for instance studying the effect of the medication of such disease. One may start from the available treated cases and then follow their disease history retrospectively for related symptoms and their healing process (Wikipedia, 2012).

As it has been explained above it follows that observational studies can be used as an alternative studies to be conducted when experimental studies cannot be conducted. Or one might plan an observational study as well. However due to lack of randomization, observational studies are prone to different sources of bias namely (1) Selection bias (2) treatment crossover bias and (3) time-dependent confounders(Faries & Kadziola (2010), Hernan, Brumbeck & Robis (2000), Hernan et al (2006), Rosenbaum (2002), Wikipedia (2012) and Delea et al (2011)). Selective bias is a bias introduced by the way subjects are recruited which makes the results not be representative. While confounders are risk factor for

disease of interest and also is associated with exposure of interest (Hammer, Du-Prel, & Blettner, 2009).

Treatment crossover in observational study is a tendency of patients to change treatments as recommended by their physicians. In observational studies this crossover depends on how physician evaluates patient's improvements. Further reasons for treatment crossover can be, (1) Experimental treatment is doing better than it was expected, as a result control patients are also allowed to receive that treatment. (2) Patients are not doing well in the treatment they were assigned and therefore physician has to change their treatments.

Therefore treatment crossover bias is the bias that is introduced due to the fact that subjects contribute information in all treatments of interest and sometimes they can cross to other treatments as well. Therefore it is difficult to know actual effect of a certain treatment if this bias is not controlled during estimation of treatment effect.

The objective of this study is to estimate treatment effect of Combination Therapy (Experimental treatment) for patients with multiple myeloma in three scenarios;

- Ignoring all sources of bias and estimate treatment effect
- Control selection bias due to the fact that it is observational study
- Control treatment crossover bias in estimating treatment effect

The structure of this report starts with introduction, data description, exploratory data analysis, methodology, statistical results, assessment of estimated crossover probabilities and weights, discussion and conclusion, limitation and recommendation for further research.

2. Data Description

The data set of this study comes from observational study with two survival end points namely; OS and PFS. It consists of 509 patients with multiple myeloma. These patients were followed up for time to OS and time to PFS. The follow-up time started in 2006 up to 2011, but patients were having different follow-up time because they were entering the study at different time point. The follow-up ended in 2011. Therefore censoring for no events was due to end of follow-up, dropout and lost follow-up.

Data set contained two treatments of interest namely, Mono therapy which was a standard treatment group while Combination therapy was experimental treatment group. Patients were followed after every two weeks, three weeks and monthly depending on how physician decides for a patient. *Table 1* summarises all variables in the data set, their description and coding.

Table 1: Summary of variables in the data set

Variable Name	Description	Code
Time_to_death	Time to OS	
Status_death	Patient's death status	1=event (death) and 0=no event
Time_to_PFS	Time to PFS	
Status_PFS	Patient's disease progression status	1=event (PFS) and 0=no event
Treatment	Patients initiated treatment	A=1 for Combination and A=0 for Mono Therapy
Albumin_cat	Albumine level at baseline (categorical)	1 >3.5 mg/dl, 2 ≤3.5 mg/dl and 9=Unknown
B2Microglob_cat	Beta 2 Microglobulines at baseline (categorical)	1 <3.5 mg/dl , 2=3.5- 5 mg/dl, 3 >5mg/dl, and 9=Unknown
Gender	Gender	1=Male and 2=Female
LOT	Line of treatments	2=Second line, 3=Third Line, 4=Fourth line, 5=5+line, 6=Best Supportive Care and 9=Unknown
MMStage_common	Stage of disease at baseline	1=Stage 1 , 2=Stage 2 , 3= stage 3 and 9=Unknown
creatclear_cat	Creatinine clearance (categorical)	1 <20ml/min , 2= 20 - <40 ml/min , 3=40 - <60 ml/min , 4=60 - <80 ml/min and 9 =Unknown
date_crossover	Date for crossover between treatments of interest	
date_nxtMMtrtmnt	Date of initiation other treatments course	

3. Exploratory Data Analysis

Exploratory data analysis was performed so as to get insight of the data. In this part of analysis cross tabulation, correlation and survival summaries and plots were explored. Percentage of events for both end points for all covariate were summarised in *Table 2*. The table showed that there were 58.35% male patients and 41.65% female patients. For the case of treatment groups, Mono Therapy had 134 patients and Combination Therapy had 375 patients.

In simple exploratory data analysis, number of events for both end points was explored by treatment groups. In OS, it was observed that in Mono therapy group there were 49 patients died and in Combination therapy there was 152 patents died. Whereas for PFS 102 patients got disease progression in Mono therapy while in Combination therapy 253 patients experienced disease progression.

Furthermore frequency distribution by status of each end point was also explored. For disease stage covariate results showed that 39.49% of all patients died, while 69.74% of all patients experienced disease progression during follow-up time. Many patients who were in disease stage 3, had experienced disease progression (38.7% of all patients) and it was followed by disease stage 2 (17.09% of all patients).

Another covariate that was of interest was line of treatments that patient received prior to entry into the study. In OS end point, 14.93% of all patients died in second line of treatments. While for PFS end point, 31.43% of all patients experienced disease progression in the same line of treatment.

For the case of Albumine level covariate, OS end point, 17.88% of all patients died and all these patients were in Albumine level ≤ 3.5 mg/dl. Further in the same Albumine level, 36.15% of all patients experienced disease progression.

Another important covariate was Beta 2 Microglobulines, for OS end point 22.59% of all patients died and these patients were in unknown category. Same category of Beta 2 Microglobulines had many patients who had experienced disease progression (38.31%). It followed by <3.5mg/dl category whereby 14.31% patients had disease progression.

Last but not least covariate was creatinine clearance. In this covariate about 19% of all patients died in 60 - <80 ml/min category. While 38.51% of all patients experienced disease progression in the same category during follow-up time.

Assessing disease progression, the same table showed that in all covariates majority of patient experienced progression as compared with those who did not have progression. The only exception was observed in unknown level of line of treatments covariate whereby few patients had progression as compared to those who did not get progression.

Table 2: Number (%) of patients for all baseline covariates by OS and PFS endpoints

Covariates	OS		PFS	
	Alive	*Died*	*No Progression*	*Progression*
Disease stage				
Stage1	71(13.95)	16(3.14)	41(8.06)	46(9.04)
Stage2	78(15.32)	51(10.02)	42(8.25)	87(17.09)
Stage3	147(28.88)	116(22.79)	66(12.97)	197(38.7)
Unknown	12(2.36)	18(3.54)	5(0.98)	25(4.91)
Gender				
Male	173 (33.99)	124(24.36)	82(16.11)	215(42.24)
Female	135(26.52)	77(15.13)	72(14.15)	140(27.5)
Line of treatment				
Second line	160(31.43)	76(14.93)	76(14.93)	160(31.43)
Third Line	76(14.93)	75(14.73)	34(6.68)	117(22.99)
Fourth Line	24(4.72)	28(5.5)	11(2.16)	41(8.06)
5+ Line	16(3.14)	11(2.16)	7(1.38)	20(3.93)
Best Supportive Care	6(1.18)	9(1.77)	3(0.59)	12(2.36)
Unknown	26(5.11)	2(0.39)	23(4.52)	5(0.98)
Albumine Level				
>3.5 mg/dl	190(37.33)	91(17.88)	97(19.06)	184(36.15)
<=3.5 mg/dl	81(15.91)	85(16.7)	41(8.06)	125(24.56)
Unknown	37(7.27)	25(4.91)	16(3.14)	46(9.04)
Beta 2 Microglobulines				
<3.5 mg/dl	98(19.25)	33(6.48)	58(11.39)	73(14.34)
3.5 - 5 mg/dl	27(5.3)	12(2.36)	14(2.75)	25(4.91)
>5 mg/dl	43(8.45)	41(8.06)	22(4.32)	62(12.18)
Unknown	140(27.5)	115(22.59)	60(11.79)	195(38.31)
Creatinine Clearance				
<20 ml/min	13(2.55)	19(3.73)	6(1.18)	26(5.11)
20 - 40 ml/min	12(2.36)	22(4.32)	5(0.98)	29(5.7)
40 - <60 ml/min	74(14.54)	52(10.22)	39(7.66)	87(17.09)
60 - <80 ml/min	199(39.1)	97(19.06)	100(19.65)	196(38.51)
Unknown	10(1.96)	11(2.16)	4(0.79)	17(3.34)

Exploring pairwise correlation between covariates to be considered in the analysis, Spearman correlation coefficient was summarized in *Table 3*. It has been observed that gender was negatively correlated with disease stage of the patients. Further negative correlation was observed between gender and Albumine level. All covariates were negatively correlation with

Creatinine clearance covariates except for Albumine. However the correlation matrix did not reveal any serious pairwise correlation between covariates.

Table 3: Spearman Correlation Coefficients for all covariates

Covariates	*Disease stage*	*Gender*	*LOT*	*Albumin*	*B2Microglob*	*creatclear*
Disease stage	1	-0.08833	-0.01228	0.08589	0.06511	-0.00167
Gender		1	-0.01871	-0.05365	0.00451	-0.03808
LOT			1	0.02239	0.04312	-0.04768
Albumin				1	0.18814	0.04284
B2Microglob					1	-0.13903
creatclear						1

4. Methodology

In this section methods that are used to estimate treatment effect in observational studies with treatment crossover were reviewed. Few models that were applicable to this project were applied to estimate treatment effects as per objectives. Before proceeding with reviews and model formulation lets describe a process of data transformation that was done in this project in order to fit weighted Cox PH models.

The original data set was survival data structure with one observation per subject. Due to nature of the weights to be estimated (Weights are time specific) standard software for fitting Cox model cannot be used for this type of analysis. In order to fit weighted Cox PH model with this type of weights, survival data structure needs to be transformed into panel (Longitudinal) data (Hernan, Brumbeck & Robis (2000) and Delea et al (2011)).

In this project data transformation was done based on observed time to event. That time to event was recorded in days and therefore it was divided into six months (180 days) interval for both end points. However due to computational difficulties (presence of quasi complete separation in some time interval) due to the fact that there was no event in some time intervals especially last time intervals. The Last 4 time intervals were combined into one time interval. Therefore last time intervals for both end points were longer than others (contains two years). In all estimation, time interval estimation was done based on binary time points.

Furthermore for the case of Status_death and Status_PFS in this transformation, for subjects with events this event was pushed to last time interval (the assumption here is that all events were observed in last time intervals), for instance [0,0,0,0,0,1]. For those who were censored for these end points were having zeros in all time intervals [0,0,0,0,0,0].

Treatment history variable, when subjects changed treatment this variable was also changing. Therefore for this case this was a vector of all treatments that was used by patients in all time intervals. Whereas for treatment crossover indicator variable those subjects who crossed treatment were censored at that time interval of which they crossed treatment. For instance [1,1,1,0,0,0] or [1,0,0,0,0,0] or [1,1,1,1,1,1] or [1,1,1,1,1,0]. Baseline covariates remained at they were by only repeating the same values in all time intervals. All statistical analysis was done in SAS version 9.2 and plots were done in R version 2.14.2 and one plot in Excel. Level of significance used was 5%.

4.1. Model Formulation of Ordinary Cox PH

Before adjusting for selection bias and crossover bias, two Cox PH models were fitted for the purpose of comparison with models that take into account complications of this project. First

model was Cox PH with only baseline treatment and second model was Cox PH with baseline treatment and baseline covariates. These models were formulated as follows:

Cox PH model with baseline treatment only

$$h_i(t \mid A_i) = h_0(t)\exp(\gamma A_i)$$

Cox PH model with baseline treatment and baseline covariates

$$h_i(t \mid A_i, V_i) = h_0(t)\exp(\gamma A_i + \beta V_i)$$

Where: β : Is the vector of unknown parameters for baseline covariates to estimated.

γ : Is unknown parameter of treatment effect to be estimated.

V_i :Is a vector of baseline covariates.

A_i : Is an indicator variable for baseline treatment.

$h_0(t)$: Is the baseline hazard at time t.

$h_i(t \mid A_i, V_i)$: Is the hazard function of subject i .

4.2. Review of Methods that are Applicable in Observational Studies

The overt bias is the bias that can be seen in observational study. This type of bias can be controlled using stratification or marching based on the covariate that make control group and treated group to differ (Rosenbaum, 2002). By doing so, people with similar characteristics in treated and control group will become comparable. There exist different methodologies for estimating treatment effect in observational studies where there is overt bias and hidden bias. Rosenbaum (2002) discussed on methods that can be applied in observational studies when there is overt bias. First: the use of the stratification on observed covariates which ensures treated and control subjects belong in the same categories. Second method is matching on observed covariates (see Rosenbaum (2002) for details).

Another popular method for dealing with selection bias in observational studies is propensity score method. Rosenbaum (2002) defined propensity score as the conditional probability of receiving treatment given the observed covariates. It can be seen in this definition this score explains the probability of being in the experimental treatment arm given all observed covariates at baseline (or prognostic factors at baseline). This score is used as a balancing score that helps the distribution of observed covariates to be similar between treatment groups that are compared.

After estimating this scores, there various methods that can be applied so that the treatment effect can be estimated in observational data. Austin (2011) discussed different methods of

using propensity score as a way of reducing bias in observational studies. Methods discussed were propensity score matching, stratification on the propensity score, IPTW using propensity score and covariate adjustment using propensity score.

The assumption of this propensity score is that there is no hidden bias in the data set, in other word bias dealt with this method is the one caused by imbalance of baseline characteristics. Therefore this method tends to balance based on observed covariates but in reality there might be imbalances due to unobserved covariates.

4.2.1. Model Formulation for Propensity Score

Literatures showed that for observational studies propensity score is unknown. Therefore it is normally estimated from the data by fitting Logistic regression with dependent variable being treatment indicator regressed on observed baseline covariates and/or prognostic factors (for details see Rosenbaum (2002) and Austin (2011)). In this project there were two treatments and therefore a logistic regression for binary (treated=1 and control=0) was fitted as follows;

$$logit[\pi(x)\,|\,V_i] = \alpha + \beta V_i$$

Where: $\pi(x) = prob(Treat = 1)$: Is the probability of receiving treatment (Combination).

β : Is unknown vector of parameters to be estimated.

4.2.1.1. Cox PH Adjusting for Stratified Propensity Score

Having estimated propensity scores as predicted probabilities of the above model, the estimated propensity scores were stratified. Rosenbaum and Rubin (1984) explained that stratifying on the quintiles of the propensity score eliminates approximately 90% of the bias due to the measured confounders when estimating a linear treatment effect. Therefore for this project estimated propensity score was stratified in five strata, each stratum was containing about 100 patients. The assumption here is that within stratum patients have similar propensity scores and therefore if propensity scores were correctly estimated, in each stratum there will balance of distribution of observed baseline covariates between experimental treatment group and control group.

That will allow estimating treatment effect by fitting a regression model with survival outcome as dependent variable and treatment indicator variable as a covariate, adjusting for stratified propensity scores. In this project Cox PH model was fitted in order to estimate treatment effect of combination therapy. The fitted model was formulated as follows:

$$h_i(t\,|\,A_i, S_i) = h_0(t)\exp(\gamma A_i + \beta S_i)$$

Where: β : Is unknown parameter for stratified propensity score.

γ : Is unknown parameter of treatment effect to be estimated.

S_i : Is the covariate that contains stratified propensity scores for subject i .

A_i : Is an indicator variable for baseline treatment for subject i .

4.2.1.2. Weighted Cox PH model with IPTW based on Propensity Score

Another approach that is commonly used to estimate treatment effect in observational studies is IPTW using estimated propensity scores. In this approach, weights for all patients were calculated as an inverse of the probability of receiving the treatment that they actually received (Lanehart, et al., 2012). In this procedure those patients who received experimental treatment will receive weight which equals to the inverse of estimated propensity scores. Those who received standard treatment their weights were equals to inverse of the 1 minus estimated propensity score.

After estimating these weights, a weighted Cox PH model with response being death status and disease progression status was regressed on baseline treatment and observed baseline covariates. This model was fitted on transformed data. Austin (2011) discussed that parameter estimate that are obtained in this model have causal interpretation which is part of the MSM family. The weighted Cox PH model fitted was formulated as follows:

$$logit[\pi(D_i(t)=1) \mid A_i, V_i] = \alpha_0 + \tau A_i + \beta V_i$$

Where: $D_i(t)$: Is a vector of patient status for all time points, (Event=1 and No event=0).

α_0 : Is intercept.

τ : Is the parameter estimate for baseline treatment effect.

β : Is the vector of parameter estimates for baseline covariates.

4.3. Review of Methods for Handling Treatment Crossover Bias

Despite of bias that comes from the nature of data, this data set had also additional complexity that patients were crossing treatment arms based on physician recommendation. Contrasting experimental studies, in observational studies people do not necessarily follow prescribed treatment regime (Hernan, Lanoy, Costagliola, & Robins, 2006). As a result conventional statistical methods of estimating treatment effect in this situation will produce

biased estimate of treatment effect due to confounding that comes from crossover. Therefore one must consider methods of estimation that take into account this complexity.

For this kind of complication, there exist different methods for estimating treatment effect in presence of treatment crossover, namely (1) excluding patients who crossed treatments (2) inclusion of crossover as a covariate in statistical model (3) Censoring patients at the time of crossover (4) In observational studies with crossover MSM using IPTW and IPCW and Nested NSM with g-estimation are normally applied.

Literature showed that the applications of MSM are often used for treatment effect estimation in observational studies with treatment crossover. Hernan et al (2006) explained possible solutions for dealing with time-varying treatment in observational studies. One of the solutions was modelling causal effects which were MSM and NSM. Marginal Structural Cox PH model using IPTW and IPCW are causal models and popular in literature, whereby many researches have been done using the model.

In general, the idea of IPTW and IPCW of estimating causal effect is that for subjects with similar baseline characteristics and they didn't drop or cross treatment, the IPTW and IPCW methods assigns bigger weights to "re-create" the population that would have been observed without crossover (Delea, Duh, Wei, & Robins, 2011). Assumptions for weighted Cox PH with both IPTW and IPCW assume no unmeasured baseline confounders as well as no model misspecification.

In this project major focus was based on these two methods IPTW and IPCW and simple method like censoring at the time of crossover was performed for comparison purpose only.

Before weighted analysis, un-weighted logistic regression was fitted on transformed data for the purpose of comparison with ordinary Cox PH models that does not take into account selection bias and crossover bias.

4.3.1. Model Formulation for Un-weighted Logistic Regression Models

In this analysis, two un-weighted logistic regression were fitted. First model was fitted with response being event (event=1 and no event=0) which was regressed on baseline treatment and baseline covariates. Second model was fitted on treatment history and baseline covariates. These two models were formulated as follows:

Un-weighted logistic regression with baseline treatment

$$logit[\pi(D_i(t)=1) \mid A_i, V_i] = \alpha_0 + \tau A_i + \beta V_i$$

Where: $D_i(t)$: Is a vector of patient status for all time points, (Event=1 and No event=0)

α_0: Is intercept.

τ : Is the parameter estimate for baseline treatment effect.

β : Is the vector of parameter estimates for baseline covariates.

Un-weighted logistic regression with treatment history

$$logit[\pi(D_i(t)=1) \mid A_i(t), V_i] = \alpha_0 + \tau A_i(t) + \beta V_i$$

Where: $A_i(t)$: Is a vector of treatment history used for all time points.

4.3.2. Cox PH model for Censoring at the Time of Crossover

As it has been introduced previously, this approach censors subjects at the time of treatment crossover. In this project there are two types of treatment crossover, first crossover between treatment of interest and second crossover to other treatments. Censoring due to treatment crossover was a minimum time to crossover between time of crossover between treatments of interest and time to crossover to other treatments. Two types censoring variables were created namely (1) artificial censoring indicator that shows whether subject crossed treatment or not (2) censoring indicator that contained all types of censoring as well as crossover censoring. In this case here those who had event were those who stayed in their treatments assigned at baseline. Moreover for those who crossed after observing an event especially for PFS endpoint were also considered that they didn't cross for that endpoint. In this approach censoring is not independent because it was linked to reasons for crossover.

Therefore censoring method applied here uses second type of censoring indicator for both end points. This analysis was done in the original survival data set because at this stage there is no need to transform the data set. The fitted Cox PH model was formulated as follows:

$$h_i(t \mid A_i, V_i) = h_0(t)\exp(\gamma A_i + \beta V_i)$$

4.3.3. Weighted Cox PH with IPTW

Marginal structural Cox PH model using IPTW method was introduced by Hernan et al (2000). They explained that traditional methods for estimating treatment effect have biased estimates for causal inference in the presence of time-dependent confounders. But Marginal Structural Cox PH model has unbiased estimate for causal inference even in the presence of time-dependent confounders and selection bias.

Hernan, et al (2000) further explained weighted Cox PH model using IPTW for estimating treatment effect. In this method weights are estimated using 4 logistic regressions with observed baseline covariates and time-dependent confounders. First two models are fitted by

assuming there is no censoring; whereby weight 1 is estimated as a ratio of two models. For numerator treatment history is regressed on previous treatment and observed baseline covariates. For denominator the treatment history is regressed on the observed baseline covariates, previous treatment and time-dependent confounders. While for weight 2, logistic regression models are fitted by assuming there is censoring. Here Censoring indicator is the response. Therefore weight 2 is a ratio of two models with censoring indicator as response regressed on baseline covariates and previous treatment for numerator weight. For denominator censoring indicator is regressed on baseline, previous treatment and time-dependent covariates. In all these models predicted probabilities are estimated and used for estimating weight 1 and 2. Therefore final weight was the product of weight 1 and 2.

Moreover Hernan, et al (2000) also distinguished between non-stabilized weights and stabilized weights. Non-stabilized weight is the one that replace numerator predicted values by 1 in both weights whereas stabilized weight is the one that has numerator predicted values. Further they continued explaining that both weights have consistent causal estimate for treatment effect but the stabilized weight has narrower confidence interval as compared to its counterpart.

The application of this IPTW and IPCW methods of estimation to data set at hand was done on the transformed panel data. However the data set had no observed time-dependent covariate. Due to that reason it restricted the direct application of the IPTW and IPCW with stabilized weights, because the direct application of stabilized weights will produce weights equal to 1 for all patients. Therefore in this analysis weights that was applied was non-stabilized.

4.3.3.1. Model Formulation for Weighted Cox PH Model with IPTW

In this analysis the probabilities are estimated by fitting two models; (1) logistic regression model with treatment history regressed on baseline covariates and previous treatment and (2) logistic regression for probability of remaining uncensored whereby censoring indicator (The one contains all types of censoring) was regressed on baseline covariates and previous received treatments. Therefore weights were estimated as an inverse of the estimated probabilities. The logistic models used for estimating weights from these probabilities were formulated as follows:

Model 1: Assuming there is no censoring in the data set.

$$logit[\pi(A_i(t)=1) \mid A_i(t-1), V_i] = \alpha + \beta_1 A_i(t-1) + \beta_2 V_i$$

Model 2: Assuming there is censoring due to crossover, dropout, lost follow-up and administrative censoring

$$logit[\pi(c_i(t)=1) \mid A_i(t-1), V_i] = \alpha + \beta_1 A_i(t-1) + \beta_2 V_i$$

After estimating weights, the weighted Cox PH model was formulated as follows;

$$logit[\pi(D_i(t)=1) \mid A_i(t), V_i] = \alpha_0 + \tau A_i(t) + \beta V_i$$

Where: $D_i(t)$: Is a vector of patient status for all time points, (Event=1 and No event=0)

$A_i(t)$: Is a vector of treatment history used for all time intervals.

V_i: Is a vector of all baseline covariates.

α_0: Is intercept.

τ: Is the parameter estimate for treatment effect.

β: Is the vector of parameter estimates for baseline covariates.

4.3.4. Weighted Cox PH with IPCW

Hernan, et al (2006) also explained another approach for estimating weight using IPCW. This method creates a scenario of missing follow-up data by censoring the follow-up of each subject at the time of crossover (Delea, Duh, Wei, & Robins, 2011). For this method subject who stayed in their initially assigned treatments will receive weight greater than 1 if there exist another subject with similar baseline/prognostic factors and who crossed treatment. Probabilities for estimating weights were estimated by fitting logistic regression with dependent variable being censoring indicator for treatment crossover (Delea, Duh, Wei, & Robins, 2011). However this method also requires time-dependent covariates in estimating weights.

For IPCW method weight 1 is estimated as a ratio of two regression models. For numerator weight, artificial censoring due to crossover was regressed on baseline treatments, baseline covariates and normal censoring indicator. While for denominator weight, regression has additional variable called time-dependent. Weight 2 is estimated based on two models. For numerator weight, normal censoring indicator was regressed on baseline treatments, baseline covariates and artificial censoring indicator for crossover. While for denominator, additional time-dependent variable is required (Hernan, Lanoy, Costagliola, & Robins, 2006). Therefore final weight is the product of weight 1 and 2.

4.3.4.1. Model Formulation of Weighted Cox PH with IPCW

The weighted Cox PH used here is the same as the one used in IPTW method. Further in this analysis non-stabilized IPCW were also used. Therefore logistic models for estimating weights were formulated as follows;

Model 1: Estimating probabilities to remain uncensored for artificial censoring ($N(t)$)

$$logit[\pi(N(t)=1)\,|\,c_i(t),A_i,V_i]=\alpha+\beta_1 c_i(t)+\beta_2 A_i+\beta_3 V_i$$

Model 2: Estimating probabilities to remain uncensored for a normal censoring ($c(t)$)

$$logit[\pi(c(t)=1)\,|\,N_i(t),A_i,V_i]=\alpha+\beta_1 N_i(t)+\beta_2 A_i+\beta_3 V_i$$

Where: $c(t)$: Is the censoring indicator that contains all types of censoring.

$N(t)$: Is the artificial censoring due to treatment crossover.

$\beta^{'}s$ are parameter estimates for treatment and baseline covariates.

4.3.5. Weighted Cox PH Model with Crossover Probabilities

The final method for estimating weights uses inverse crossover probabilities. For this method probabilities of crossing were estimated in different assumptions for treatment crossover. Models used for estimating crossover probabilities were formulated as follows:

Prob 1: The assumption is that crossover probability is independent of treatment received.

$$logit[\pi(N(t)=0)\,|\,V_i]=\alpha+\beta V_i$$

Prob 2: The assumption is that crossover probability depends on the baseline treatment and baseline covariates

$$logit[\pi(N(t)=0)\,|\,A_i,V_i]=\alpha+\gamma A_i+\beta V_i$$

Prob 3: The assumption is that crossover probability depends on treatment history and baseline covariates

$$logit[\pi(N(t)=0)\,|\,A_i(t),V_i]=\alpha+\lambda A_i(t)+\beta V_i$$

Prob 4: The assumption is that crossover probability depends on the previous treatment and baseline covariates

$$logit[\pi(N(t)=0)\,|\,A_i(t-1),V_i]=\alpha+\tau A_i(t-1)+\beta V_i$$

Where: β : Is the vector of unknown parameter estimates for baseline covariates

γ , λ and τ : Are parameter estimate for treatment in different crossover assumptions.

$\pi(N(t)=0)$: Is the probability of crossing treatments.

The weighted Cox PH model was fitted for each estimated weight. The fitted model was the same as the one used in IPTW and IPCW methods.

5. Statistical Results

5.1. Exploring Crossover Patterns

Assessing the magnitude of treatment crossover (between treatments of interest and other treatments), crossover pattern was explored for all types of crossover that exist in the data set. It was observed that in this study patients were crossing treatments in three ways, first: patients were crossing between treatments of interest (Mono and Combination therapy), second: patients were crossing from either Mono or Combination therapy direct to other treatments and third: patients were crossing twice, first between treatments of interest and later they were crossing to other treatments.

In general there was 375 (73.67%) patients switched from their initially assigned treatment to either Mono therapy or Combination therapy and then to other treatments or direct to other treatments or even crossing between treatments of interest only. *Table 4* summarises all types of crossover for all patients with their initially assigned treatment groups and median time for treatment crossover. Results showed that many patients crossed from treatments of interest direct to other treatments (52 patients in Mono therapy and 131 patients for Combination therapy). Another crossover was for those patients who were crossing twice, for Mono therapy there were 40 patients whereas for Combination therapy there were 101 patients. There were very few patients crossing between treatments of interest only that is 8 patients in Mono therapy and 43 patients in Combination therapy.

In general 144 patients who were initially assigned to Combination therapy crossed to Mono therapy and 48 patients who were initially assigned in Mono Therapy crossed to Combination therapy. Patients who crossed to other treatments from Mono therapy were 92 and those who crossed to other treatments from Combination therapy were 232 patients.

Moreover median time for cross over between treatments of interest was 42 days while the median time for crossover to other treatments was 210 days. Being explored both types of crossover, it has been observed that patients were crossing between experimental treatments quickly (Mono therapy median time for crossover was 51 days and Combination therapy median time for crossover was 31.5 days) as it was compared to time used to cross to other treatments (median time for crossing from Mono therapy was 243 days and for Combination therapy was 201 days).

Table 4: Description of crossover pattern by treatments groups

Types of Crossover	Mono Therapy	Combination Therapy
Crossover between treatments of interest and then to other treatments	40	101
Stayed on treatment	34	100
Crossover from treatments of interest direct to others treatments	52	131
Crossover between treatments of interest	8	43
Total Patients	134	375
Crossover to Combination therapy	48	
Crossover to Mono therapy		144
Crossover from Mono to Other Treatment	92	
Crossover from Combination to Other Treatments	232	
Median Time to crossover between treatments of interest	42 Days	
Median Time to crossover other treatments	210 Days	
Median time for Crossover to either Combination or Mono therapy	51 Days	31.5 Days
Median time for Crossover from either Combination or Mono Therapy to Other Treatments	243 Days	201 Days

Further, it was observed that for Mono therapy 100(74.6%) of all patients in the study crosses to either treatments of interest or to other treatments or crossed twice while for Combination therapy 275 (73.3%) also crossed treatments. Therefore during the study period for both Mono therapy and Combination therapy only few patients stayed in their initially assigned treatments up to the end of follow up. *Figure 1* summarises all types of crossover as compared to those who stayed.

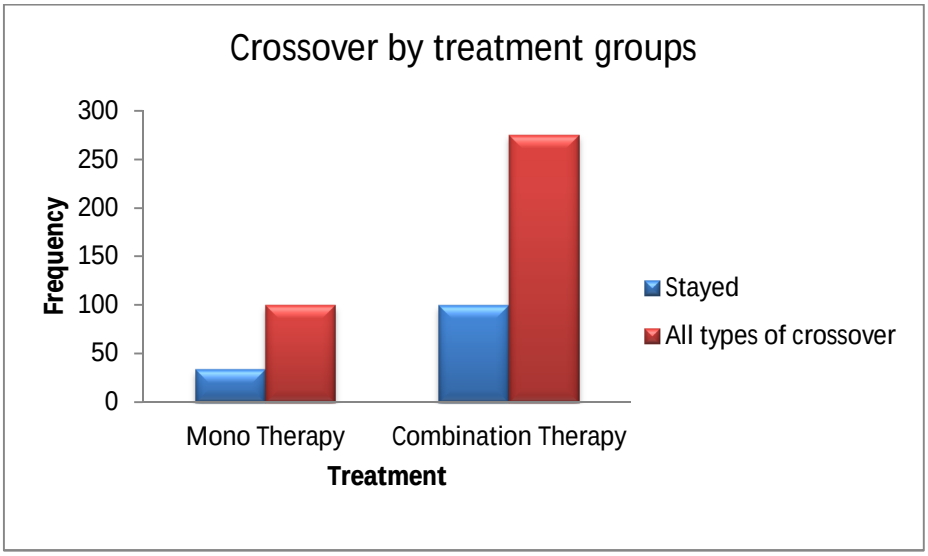

Figure 1: A plot of all types of treatment crossover by their initially assigned treatment groups

Exploratory data analysis on crossover by disease stage and line of treatment was also performed and summarised in *Table 5*. This was done because disease stage shows how severe the patient was and for line of treatments showed the more lines of treatments the more severe the patient was. This gave an idea what type of patients was crossing treatment. For disease stage at baseline the crossover between treatments of interest and/or to other treatments was higher in all disease stages, but there were many patients (39.69%) crossed who were in disease stage 3 as compared to other disease stages. It was followed by stage 2 (18.47% of all patients). Moreover *Table 5* showed that for the case of line of treatment at baseline many patients were crossing either between treatments of interest and/or other treatments in all categories of line of treatments. But majority of patients in second lines of treatments (34.58%) crossed treatment.

Table 5: Count (%) of crossover by disease stage and lines of treatments

Covariates	*Crossover*		*Total*
	Yes	*No*	
Disease stage			
Stage 1	61(11.98)	26(5.11)	87(17.09)
Stage 2	94(18.47)	35(6.88)	129(25.34)
Stage 3	202(39.69)	61(11.98)	263(51.67)
Unknown	18(3.54)	12(2.36)	30(5.89)
Line of Treatment			
Second Line	176(34.58)	60(11.79)	236(46.37)
Third Line	115(22.59)	36(7.07)	151(29.67)
Fourth Line	35(6.88)	17(3.34)	52(10.22)
5+ Line	20(3.93)	7(1.38)	27(5.3)
Best Supportive Care	9(1.77)	6(1.18)	15(2.95)
Unknown	20(3.93)	8(1.57)	28(5.5)
Total	375(73.67)	134(26.33)	509

Moreover, Kaplan-Meier plots by treatment groups for both end points were also explored. *Figure 2* showed that in both end points survival curves were touching and even crossing, especially for PFS plot whereby curves crossed in early time of follow-up. Moreover these plots also showed lack of proportionality in estimated curves, which suggests violation of proportional hazards assumption. As a result all conventional survival analysis such as Logrank test for comparing survival curves and Cox proportion model are no longer appropriate for this type of data set. In these Kaplan-Meier plots it should be noted that patients who crossed treatment contribute information in both curves.

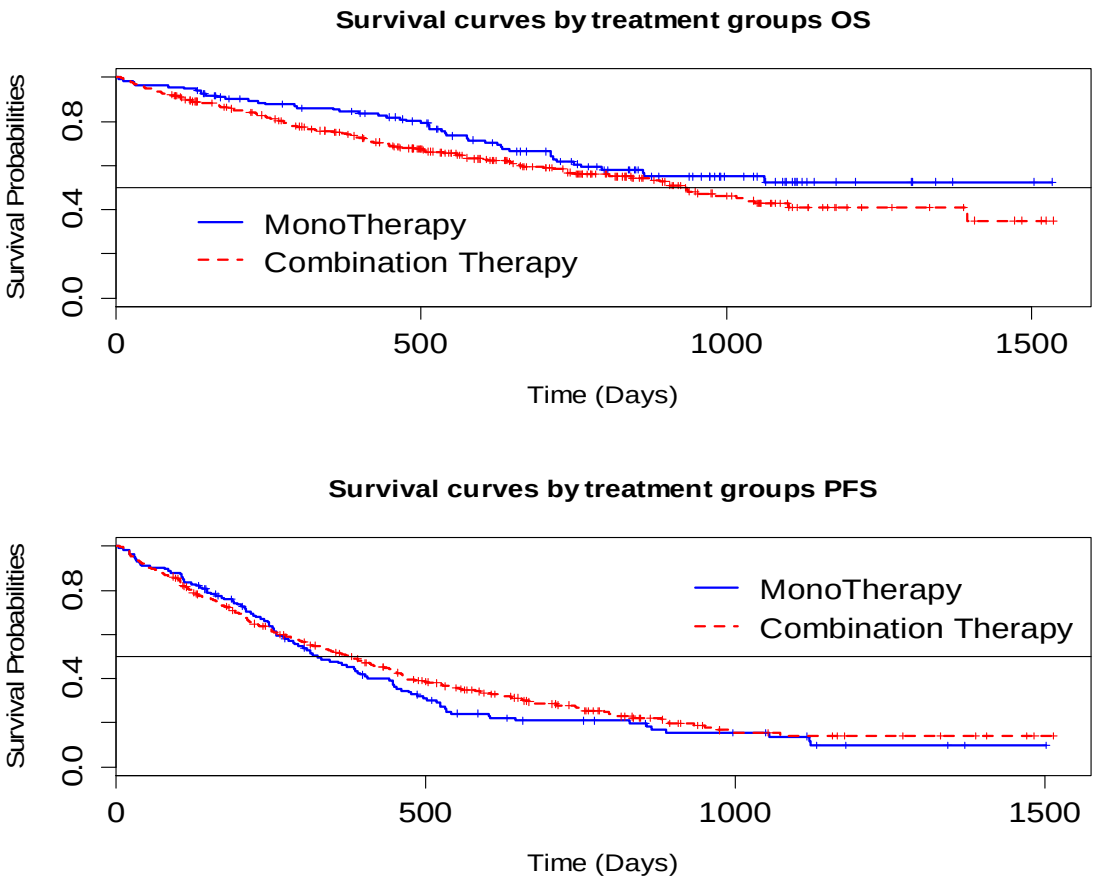

Figure 2: Kaplan-Meier Survival curves plots for the original data, horizontal line represents median survival probability

5.2. Statistical Analysis

5.2.1. Models Fitted for OS End Point

Ordinary Cox PH models: Results for OS end point showed these models had positive parameter estimates which were supporting standard treatment. The difference between estimates for model with only baseline treatment and the one with baseline treatment and baseline covariates was small. But Confidence interval for model with baseline treatment and baseline covariates was a bit wider as compared to its counterpart.

Propensity score methods: Estimates that were based on propensity score methods (Cox PH with stratified propensity score and weighted Cox PH with IPTW on propensity score model) showed the estimates with same direction with previous models. Further it was observed that estimate of Cox PH model with stratified propensity score was close to ordinary Cox PH model with baseline covariates and baseline treatment. While using weighted Cox PH with

IPTW based on propensity score, estimate was smaller compared to all previous Cox PH models. In both models there was no statistical evidence for treatment effect.

Un-weighted logistic regression: In this analysis un-weighted logistic regression model (for baseline treatment and treatment history) was fitted in a transformed data set. For the case of un-weighted logistic model with baseline treatment, estimate was close to ordinary Cox PH model (0.2486). This shows these two models are estimating the same thing. For the case of un-weighted logistic model with treatment history, estimated treatment effect was -0.3843, which was now supporting combination therapy. For this model the parameter estimate was statistically significant.

Censoring method for handling crossover bias: In this analysis it has been observed that estimate for OS end points was -0.161 but its confidence interval was wider as compared to other methods of estimating treatment effect. There was not statistical evidence of treatment effect.

Weighted Cox PH with IPTW and IPCW: For these analyses non-stabilizing weights were estimated as explained previously. In OS end point results showed that estimates for both weights were in the same direction (supporting combination Therapy). However for IPCW weight estimate was more negative as compared to result of IPTW. For both methods of estimating weights, confidence intervals for parameter estimates were wider. This was expected because literatures explained that non-stabilized IPTW and IPCW produce consistent parameter estimates but with wider confidence interval as compared to stabilized weights. Different conclusion was also observed, for IPTW there was no statistical evidence of treatment effect. Whereas IPCW method, there was statistical evidence of treatment effect.

Weighed Cox PH models based on crossover probability (Prob 1 to 4): In these analyses the probability for crossing treatment were estimated in four different assumptions as explained in methodology part. Under all estimated probabilities for crossover it was observed that time was associated with probability of crossing treatment. Whereas all other baseline covariates were not associated with the probability of crossing (results not shown). Further adding treatment at baseline (model for Prob 2) crossover probability was not associated with baseline treatment. Model for Prob 3, it was observed that treatment history was associated with crossover probabilities. Finally model for Prob 4, it was also observed that previous treatment was associated with probability of treatment crossover.

Under all these assumptions, weighted Cox PH models were fitted in order to estimate treatment effect. In assumption 1 and 2 for OS end point estimate were similar, and estimates were both statistically significant. Further it was observed that as probabilities for crossing

treatment was modified (assumption 3 and 4) the parameter estimate for treatment was becoming more negative (meaning more treatment effect for Combination therapy).

Parameter estimates for treatment effect and their confidence intervals for all fitted models of OS end point are summarised in *Table 6.* Forest plot was used to plot these estimates with their confidence interval (see *Figure 3).*

Having discussed on how the parameter estimate for treatment effect was changing in different in all fitted models. Other adjusted covariates were also evaluated (see detailed results for all fitted models and parameter estimates in Appendix).

Ordinary Cox PH model adjusting for baseline covariates, disease stage and Beta 2 Microglobulines covariates were statistically significant for OS. For the case of weighted Cox PH with IPTW based on propensity score, Albumine level and disease stage were associated with OS end point.

Un-weighted logistic models (for treatment at baseline and treatment history) fitted on panel data, in these models Albumine level, disease stage and Beta 2 Microglobulines were associated with OS end point.

Moreover those models that were taking into account treatment crossover, for censoring methods only disease stage was associated with OS. Whereas weighted Cox model with IPTW and IPCW, disease stage, Albumine, Beta 2 Microglobulines and Creatinine clearance were associated with OS.

Finally for models that were fitted using weight that was estimated as inverse crossover probability time, Albumine, disease stage and Beta 2 Microglobulines were also associated with OS end point. In all fitted models there was no statistical evidence of gender effect in OS end point.

Table 6: Parameter estimates for treatment (Combination therapy) with 95% confidence limits for OS end point

Models	Estimate	*95% Confidence Limits*
Ordinary, Propensity Score and Un-weighted Models		
Ordinary Cox PH (Treatment at baseline only)	0.2963	(-0.0263, 0.619)
Ordinary Cox PH (Baseline treatment and baseline covariates)	0.2864	(-0.0566, 0.6297)
Cox PH adjusting for stratified propensity score	0.3107	(-0.0294, 0.6508)
Weighted Cox PH(IPTW on propensity scores)	0.1239	(-0.2485, 0.4964)
Un-Weighted Logistic(Treatment at baseline)	0.2486	(-0.1156, 0.6129)
Un-Weighted Logistic(Treatment history)	-0.3843	(-0.7007, -0.0678)
Censoring and Weighted Cox PH Models For Crossover		
Censoring at the time of Crossover	-0.161	(-0.803, 0.4812)
Weighted Cox PH with IPTW (Weight 1: Non-Stabilizing weight)	-0.463	(-0.9358, 0.0099)
Weighted Cox PH with IPCW (Weight 2: Non-Stabilizing weight)	-0.6971	(-1.237, -0.1572)
Weighted Cox PH Model 1 (Weight 1: Equal Probability to crossover)	-0.3908	(-0.7471, -0.0345)
Weighted Cox PH Model 2 (Weight 2: Probability to crossover depend on baseline treatment)	-0.3895	(-0.7456, -0.0334)
Weighted Cox PH Model 3 (Weight 3: Probability to crossover depend on treatment history)	-0.4111	(-0.755, -0.0673)
Weighted Cox PH Model 4 (Weight 4: Probability to crossover depend on previous treatment)	-0.4458	(-0.8152, -0.0764)

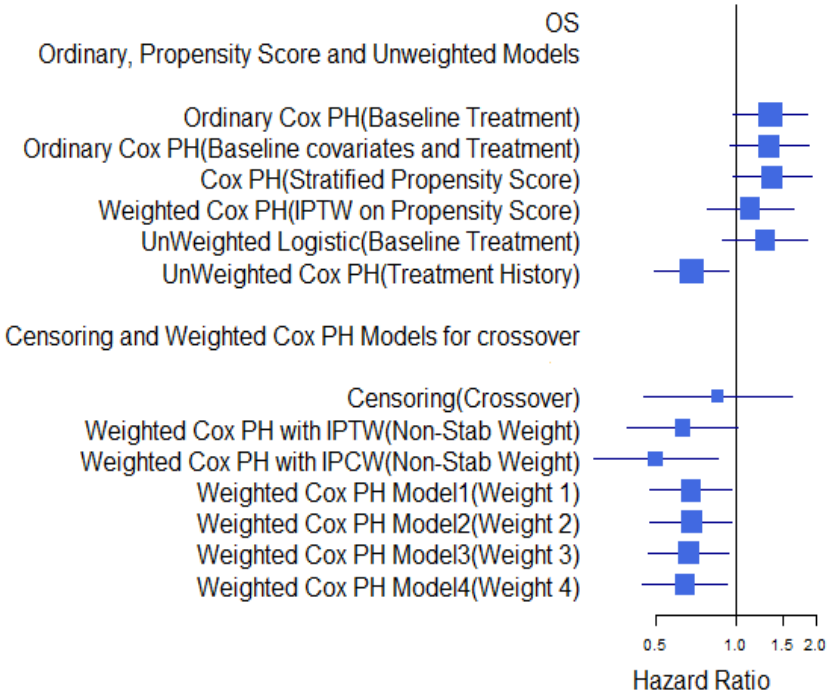

Figure 3: A plot that summarises estimates of hazard ratio of treatment effect with 95% confidence limits for all fitted models of OS end point

5.2.2. Models Fitted for PFS End Point

In this analysis, all results for PFS end point are summarized in *Table 7* and *Figure 4* summarises results in terms of plot.

Ordinary Cox PH model: This analysis showed that estimate of treatment effect was supporting combination therapy. Estimates of ordinary Cox PH with baseline treatment only was smaller than for ordinary Cox PH with baseline treatment and baseline covariates. In both models there was no statistical evidence of treatment effect.

Propensity score methods: These models showed results that were in same direction as the one obtained in ordinary Cox PH models. Estimate for Cox PH adjusting for stratified propensity scores was similar to the ordinary Cox PH model. Whereas weighted Cox PH with IPTW based on propensity score was smaller than all previous models. Its confidence interval was a little bit wider as compared to its counterpart. In both models there was no statistical evidence of treatment effect.

Un-weighted logistic regression models: In this analysis small difference was observed for parameter estimates as compared previous analysis. Even confidence intervals were similar except for un-weighted logistic regression with treatment history whereby estimated confidence interval was narrower. There was no statistical evidence of treatment effect.

Censoring Methods: Parameter estimates for this was close to the one obtained in ordinary Cox PH model with baseline treatment and baseline covariates (-0.0583). However it confidence interval was very wide compared to previous analysis. This may be explained as a result of losing a lot of information due to censoring those who crossed treatment.

Weighted Cox PH with IPTW and IPCW: Results for PFS showed that estimates were larger and for IPCW model it was even positive (0.3774) which was supporting Mono therapy. Confidence interval for IPTW was more or less narrower than IPCW. However in both models there was no statistical evidence of significant treatment effect.

Weighed Cox PH models based on probability to crossover (Prob 1 to 4): Results for prob 1 to 4, for PFS end point there was more or less similar trend except for parameter estimate under assumption 3 which had slightly larger estimate (-0.2358) as compared to other estimates in other assumptions. Estimate under assumption 4 was smaller as compared to all other assumptions.

Other adjusted covariates for PFS end point were also evaluated (see details for all fitted models and parameter estimates in Appendix). Ordinary Cox PH model adjusting for baseline covariates, disease stage and Beta 2 Microglobulines covariates were statistically significant for PFS. For the case of Weighted Cox PH with IPTW based on propensity score, disease stage and Beta 2 Microglobulines were associated with PFS end point.

Un-weighted logistic models (for treatment at baseline and treatment history) fitted on panel data, in these models disease stage, line of treatment and Beta 2 Microglobulines were associated with PFS end points. Moreover models that were taking into account treatment crossover, for censoring methods only disease stage was associated with PFS. For weighted Cox PH model with IPTW and IPCW Albumine level, disease stage and Beta 2 Microglobulines were associated with PFS.

Finally weighted models based on weight that was estimated as inverse probability of crossing treatment; time, Albumine, disease stage and Beta 2 Microglobulines were associated with PFS end point. In all fitted models for PFS there was no statistical evidence of gender effect for this end point.

Table 7: Parameter estimates for treatment (Combination Therapy) with 95% confidence limits for PFS end point

Models	Estimate	*95% Confidence Limits*
Ordinary, Propensity Score and Un-weighted Models		
Ordinary Cox PH (Treatment at baseline only)	-0.1131	(-0.3439, 0.1169)
Ordinary Cox PH (Baseline treatment and baseline covariates)	-0.0649	(-0.3065, 0.1765)
Cox PH adjusting for stratified propensity score	-0.0751	(-0.3175, 0.1672)
Weighted Cox PH(IPTW on propensity scores)	-0.1202	(-0.4007, 0.1603)
Un-Weighted Logistic(Treatment at baseline)	-0.1308	(-0.4007, 0.1392)
Un-Weighted Logistic(Treatment history)	-0.1065	(-0.3701, 0.1572)
Censoring and Weighted Cox PH Models For Crossover		
Censoring method at the time of Crossover	-0.0583	(-0.4732, 0.357)
Weighted Cox PH with IPTW (Weight 1: Non-Stabilizing weight)	-0.1468	(-0.4822, 0.1886)
Weighted Cox PH with IPCW (Weight 2: Non-Stabilizing weight)	0.3774	(-0.1335, 0.8882)
Weighted Cox PH Model 1 (Weight 1: Equal Probability to crossover)	-0.3209	(-0.6136, -0.0281)
Weighted Cox PH Model 2 (Weight 2: Probability to crossover depend on baseline treatment)	-0.2955	(-0.588, -0.003)
Weighted Cox PH Model 3 (Weight 3: Probability to crossover depend on treatment history)	-0.2358	(-0.5277, 0.0561)
Weighted Cox PH Model 4 (Weight 4: Probability to crossover depend on previous treatment)	-0.5398	(-0.8401, -0.2394)

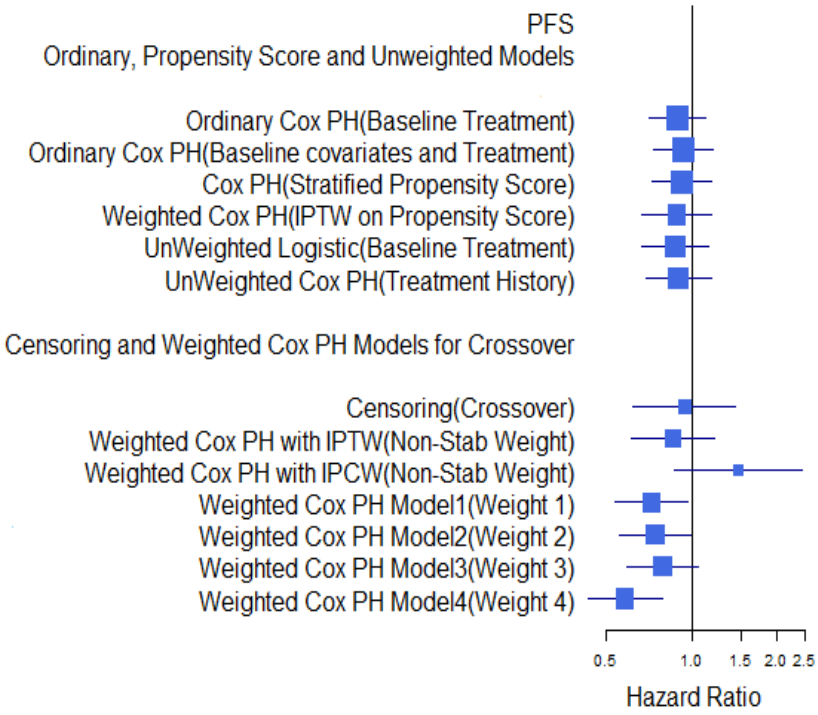

Figure 4: A plot that summarises estimates of hazard ratio of treatment effect with 95% confidence limits for all fitted models of PFS end point

6. Diagnosis of Estimated Probabilities and Weights

6.1. Assessment of the Estimated Propensity Scores by Treatment Groups

The estimated propensity scores for the analysis which uses propensity scores for balancing baseline difference treatment groups were assessed using box plot as explained by Rosenbaum and Rubin (1984). In *Figure 5*, it was observed that the distributions of estimated scores in both treatment arms were overlapping. This was an indication of balance of the baseline covariates by treatment that was assigned to patients. This makes the two treatment groups to be comparable based on the observed baseline covariates. However there were few patients in Combination therapy that were having large values of estimated propensity score and some with small values. The estimated median propensity score for patients in combination therapy were relatively higher as compared to Mono therapy median propensity scores.

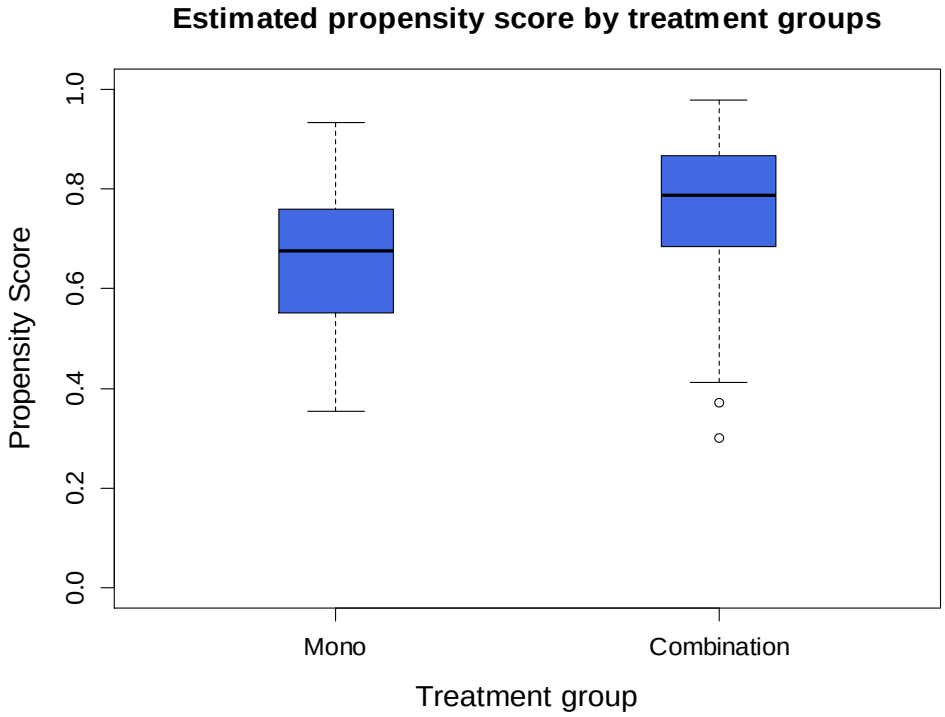

Figure 5: A boxplot for estimated propensity scores by treatment groups

6.2. Assessment of Estimated Weights for IPTW and IPCW

Weights that were estimated by IPTW and IPCW were also assessed for presence of outliers and their distributional trend. *Figure 6* summarises all these weights for all time intervals. In this plot log weights was used in order to reduce scale for the estimated weights. It was observed that there were some outlying weights for both weights. For IPCW it was also observed that as time increases weights were also increasing. This means that probability to stay in treatment that they were assigned at baseline was becoming smaller as time increases.

Weights that were estimated using IPTW cannot be interpreted directly because it had probability to be in experimental treatment and probability to remain uncensored. Moreover for weights that was estimated based on propensity scores were also assessed (Results not shown) and there were no outliers present in estimated weights.

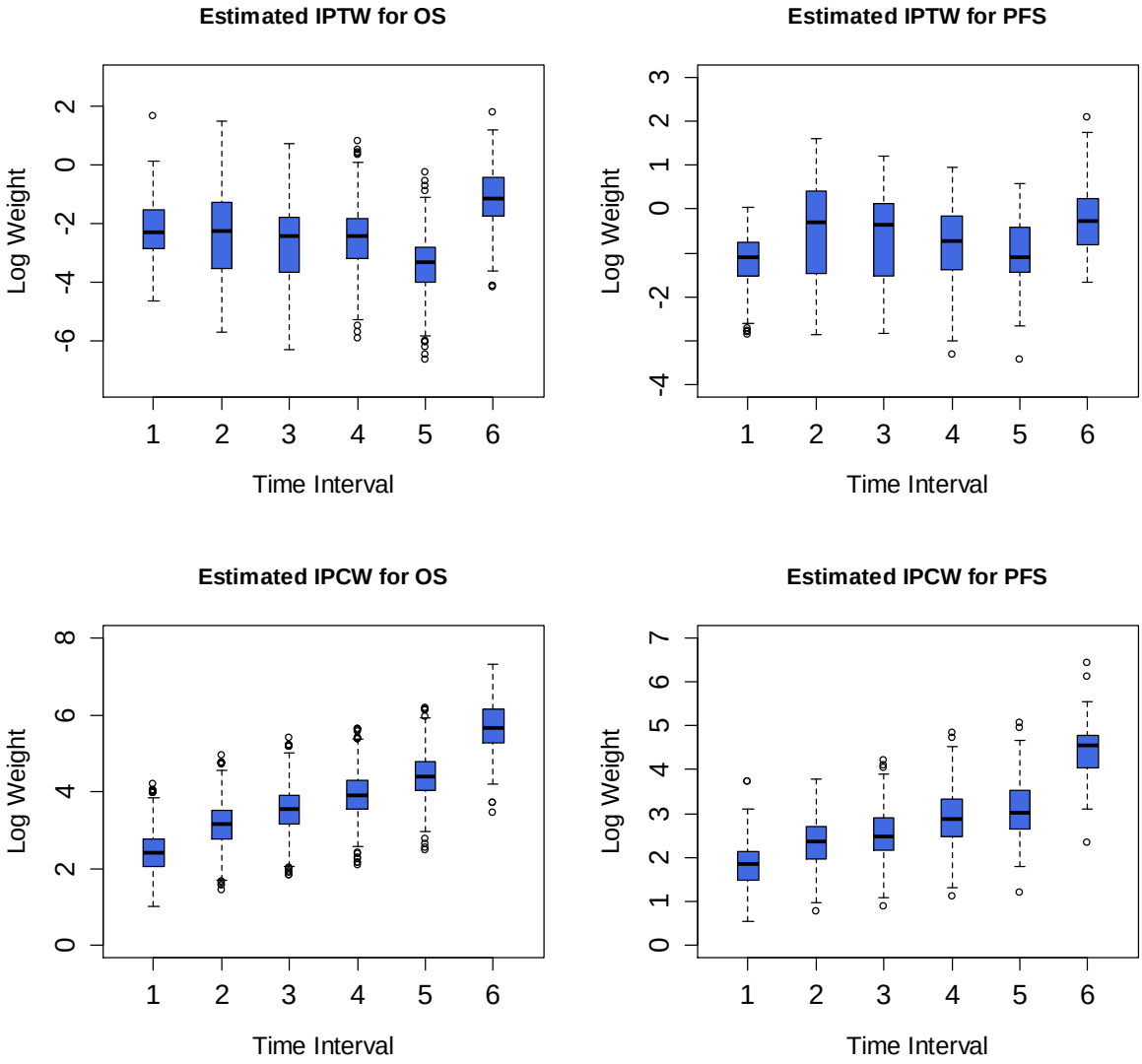

Figure 6: A plot for weights that were used in weighted Cox PH model with IPTW and IPCW

6.3. Assessment of Estimated Crossover Probabilities and its Weights in 4 Assumptions

For the case of probabilities to crossover (weighted model 1 to 4) for few selected patient that were having long follow-up time in that study were also assessed. The trend showed that there was small probability for crossing treatments at time interval 1. However the crossover probabilities increased at time interval 2 (See *Figure 7*), but as time increases the probability to crossover was increasing at lower rate. This means that those who entered in the study

earlier majority were crossing treatments at time interval 2. Reasons for crossing treatment might be treatment failure, toxicity, metastasis etc. Moreover similar trend were reflected for weights that were estimated based on these probabilities in both end points (see *Figure 8 and 9*). These plots showed big drop of weights from time interval 1 to 2 but from there weights remained constant and few outlying weights were also observed in both end points.

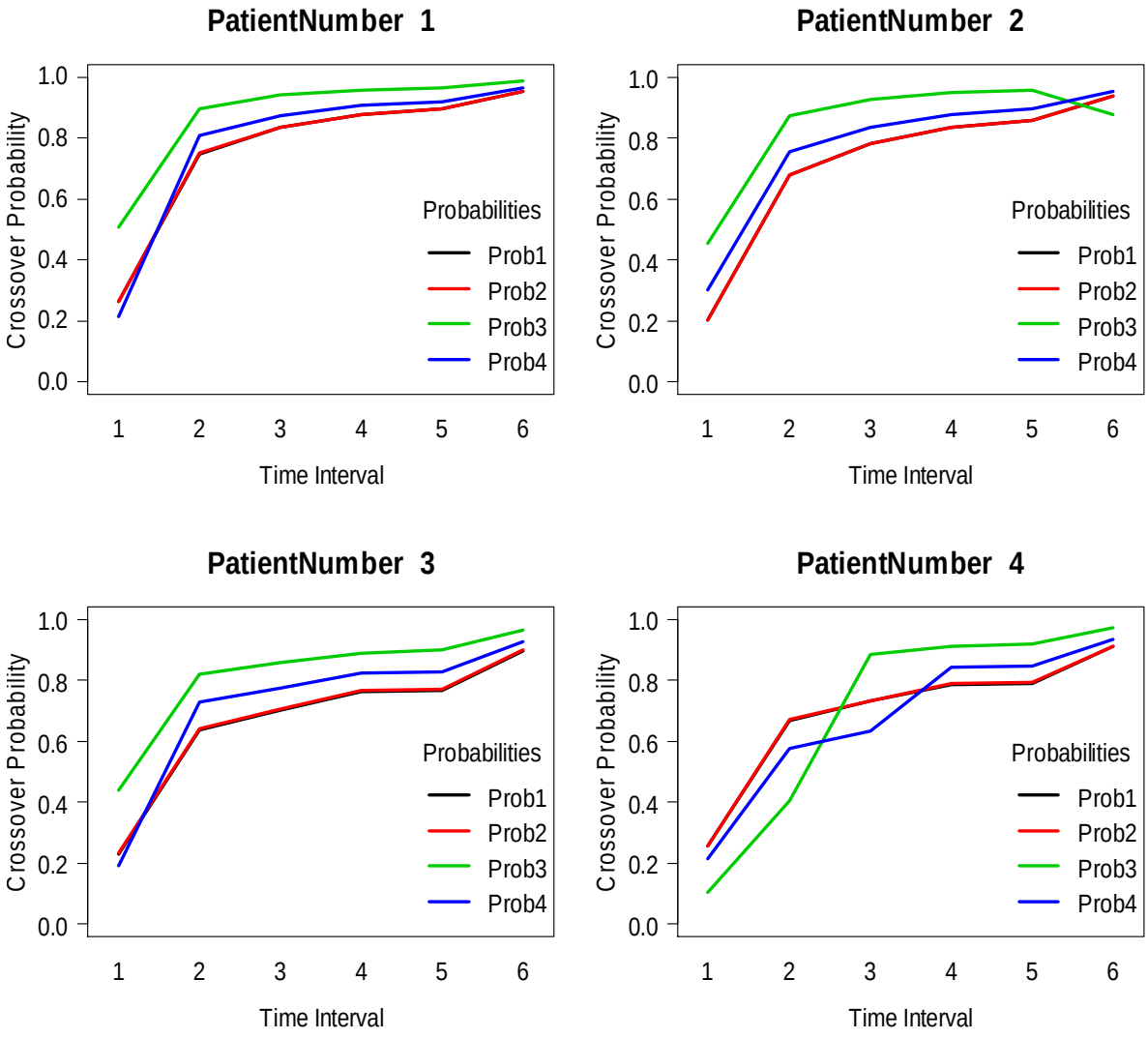

Figure 7: Estimated Probabilities to cross treatment for few selected patients, patient 1 and 2 correspond to patient ID 192 and 340 for OS and patient 3 and 4 corresponds to patient ID 222 and 451 for PFS

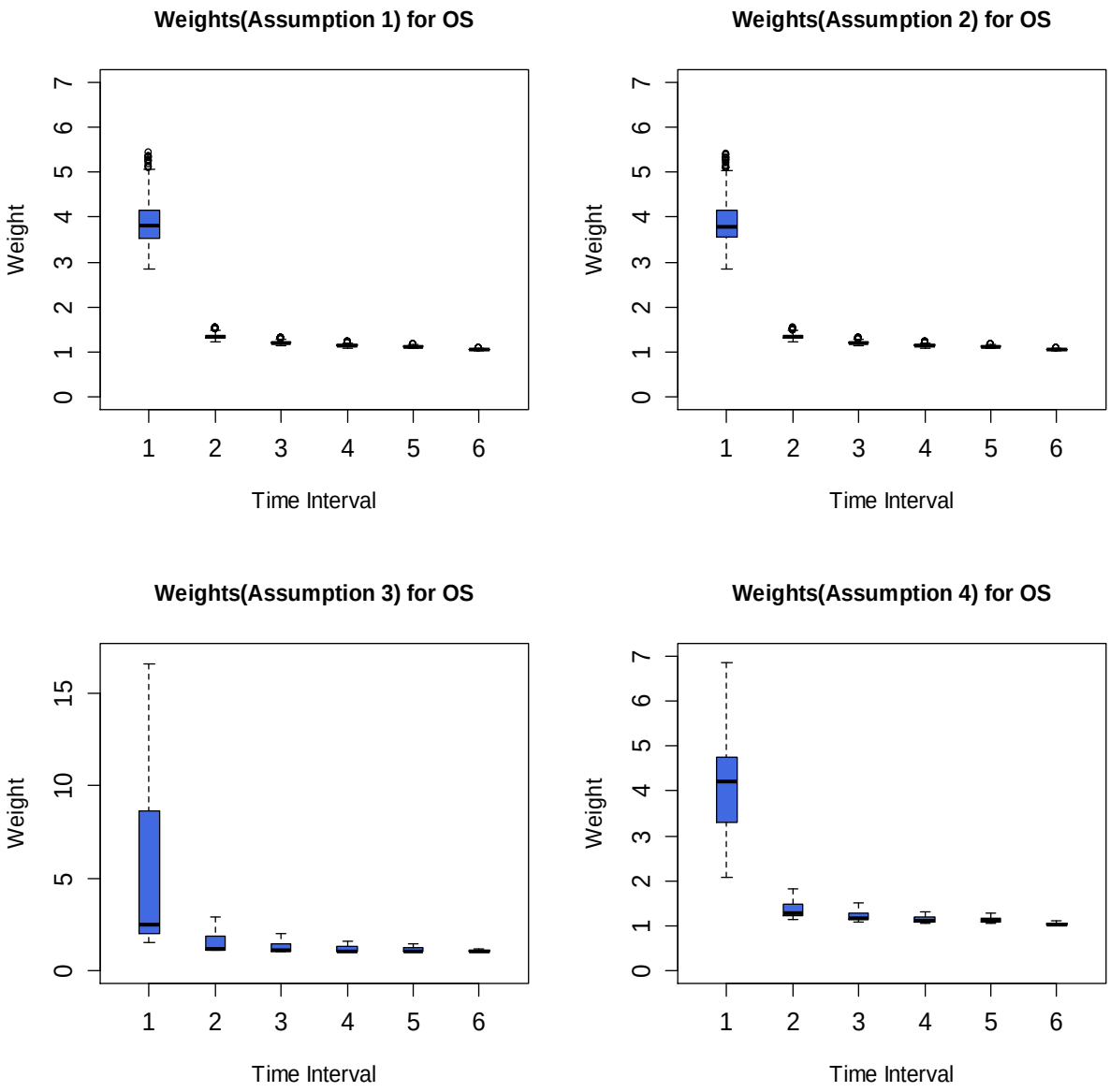

Figure 8: A plot for weights that were used in weighted Cox PH model with 4 assumptions for OS end point

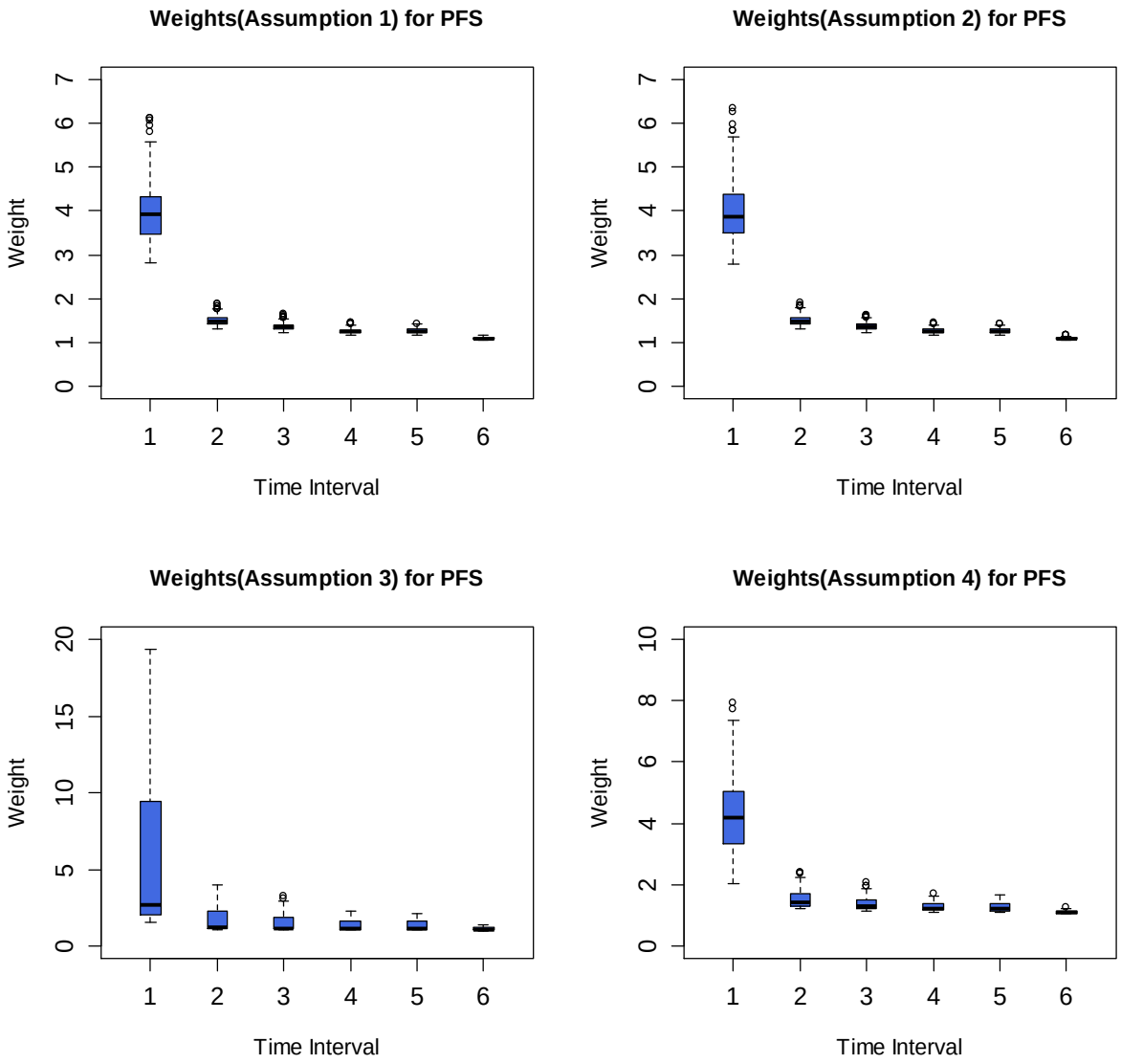

Figure 9: A plot for weights that were used in weighted Cox PH model with 4 assumptions for PFS end point

7. Discussion and Conclusion

Ordinary Cox PH model fitted here assumes that patients were randomized in treatment arms. In this analysis treatment effect that is estimated is biased because it ignores the nature study (observational study). However even if there was randomization the estimates would have been biased due to treatment crossover.

Propensity score methods have been widely applied as a way of balancing baseline difference of treated group and control group. Models fitted by using this method control selection bias in observational studies based on observed baseline covariates. The drawback of these methods is that if it happens that there are unmeasured confounders for a given end point; estimates will still be biased unless unobserved confounders are highly correlated with one of the observed (Stukel, et al., 2007). Further for IPTW based on propensity score method sometimes may produce very large weights due to extreme propensity scores as a results weights become unstable. However the advantage this method over matching on propensity scores or stratification on propensity scores is that it includes all subjects in the analysis (Lanehart, et al., 2012).

For Censoring at the time of crossover methods, in OS end points estimate was different from others. This can be explained as a result of throwing a lot of information by censoring patients that crossed treatment before experiencing events. In exploratory data analysis for crossover (*Table 4*) showed in both treatment groups there few patients that remained in their initially assigned treatment groups. Therefore by censoring all those patients that crossed treatments, the analysis will be based on only few patients. For those few patients there will be very few events especially for OS. As a result power of the test becomes very low to detect treatment effect because in survival analysis number of event is important for power of the test. Note that for PFS estimate was close ordinary Cox PH model and Cox PH adjusting for propensity scores, because here some patients with larger time to crossover than time to progression as a result for this end point they were considered that they did not cross treatments.

Weighted Cox with IPTW and IPCW models are widely used for observational data with treatment crossover. The modified version of that model was fitted using non-stabilised weights. The only weakness of this model is that using non-stabilized weight has wider confidence interval as compared to stabilized weights, but the parameter estimates are consistent. Challenge for this method is that it uses patients that stayed in modelling it's not clear for a situation where many patients crossed treatment and only few stayed whether the estimate will still be okay as compared to a situation whereby no many patients crossed.

Furthermore models that were fitted based on weights that were estimated based on the probability to crossover, these models showed similar trend of treatment effect. However as the weights were modified the treatment effect were also improving.

In all weighted Cox PH models, it has been observed that weights had larger impact for OS as compared to PFS. Investigators explained that, this was expected because patients who were crossing treatments majority had experienced disease progression already.

To summarise the discussion, it has been observed that weighted analysis restores the diluted treatment effect due to crossover and selection bias. Therefore methods that were adjusting for a problem of crossover in the analysis were preferred for this study. Furthermore Interpreting Kaplan-Meier plot in exploratory data analysis in this project is wrong because those patients who crossed treatment were contributing information of both treatment groups. As a result treatments groups cannot be compared because it's difficult to disentangle treatment effect from standard effect.

8. Limitations and Recommendations

The applications of MSM with inverse probability weighting and NSM using g-estimation with stabilized weights were limited for this project due to absence of time-dependent covariates. These covariates are the one used by physician to change treatment for patients. Therefore further research can be done if this information will be available and compare results with results obtained in this project.

Moreover the literature showed that valid confidence interval for estimates that uses MSM and NSM is the one obtained by bootstrapping. Therefore it is recommended in future to estimate confidence intervals for parameter estimate using bootstrap.

Reference

Austin, P. C. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *PMC*, 399–424.

Curtis, L. H., Hammill, B. G., Eisenstein, E. L., Kramer, J. M., & Anstrom, K. J. (2007). Using Inverse Probability-Weighted Estimators in Comparative Effectiveness Analyses with Observational Databases . S103-S107.

Delea, T. E., Duh, M. S., Wei, L.-J., & Robins, J. M. (2011, May 23). ISPOR Workshop-Session II. *Methods to adjust for selection crossover bias in survival analysis in assessment of cost-effectiveness of cancer therapies*.

Faries, D. F., & Kadziola, Z. A. (2010). Analysis of Longitudinal Observational Data Using Marginal Structural Models. In D. E. Faries, A. C. Leon, J. M. Haro, & R. L. Obenchain. SAS Institute Inc.

Hammer, G. P., Du-Prel, J.-B., & Blettner, M. (2009). Avoiding Bias in Observational Studies. *PMC*, 664–668.

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*, 561-570.

Hernan, M. A., Cole, s. R., Margolick, J., Cohen, M., & Robins, J. M. (2005). Structural accerated failure time models for survival analysis in studies with time-varying treatments. *Wiley InterScience*, 477-491.

Hernan, M. A., Lanoy, E., Costagliola, D., & Robins, J. M. (2006). Comparing Dynamic Treatment Regimes via Inverse Probability Weighting. *PubMed*, 237-242.

Lanehart, R. E., de Gil, P. R., Kim, E. S., Bellara, A. P., Kromrey, J. D., & Lee, R. S. (2012). Propensity score analysis and assessment of propensity approaches using SAS procedures. *SAS Global forum 2012*.

Rosenbaum, P. R. (2002). *Observational Studies.* Springer-Verlag New York, Inc.

Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing Bias in Observational Studies Using Subclassification on the Propensity Score. *JSTOR*, 516-524.

Stukel, T. A., Fisher, E. S., Wennberg, D. E., Alter, D. A., Gottlieb, D. J., & Vermeulen, M. J. (2007). Analysis of Observervational Studies in the Presence of Treatment Selection Bias. *JAMA*, 278-285.

Wikipedia. (2012, July 01). *Observational Study*. Retrieved from Wikipedia, The Free Encyclopedia: http://en.wikipedia.org/wiki/Observational_study

Appendix

Table 8: Estimates (se) for Ordinary Cox PH model for baseline treatment and baseline covariates

Covariate	*OS*			*PFS*		
	Estimate (se)	*P-Value*	*95% Hazard Ratio Confidence Limits*	*Estimate (se)*	*P-Value*	*95% Hazard Ratio Confidence Limits*
Baseline Treatment(Combination)	0.28644(0.17521)	0.1021	(0.945, 1.877)	-0.06489(0.12302)	0.5978	(0.736, 1.193)
Albumine						
>3.5 mg/dl	0.0068(0.27356)	0.9802	(0.589 ,1.721)	-0.0517(0.19445)	0.7903	(0.649, 1.39)
<=3.5 mg/dl	0.52642(0.27064)	0.0518	(0.996, 2.877)	0.26035(0.19755)	0.1875	(0.881, 1.911)
Gender (Female)	0.00168(0.15093)	0.9911	(0.745, 1.346)	-0.04077(0.1135)	0.7194	(0.769, 1.199)
Line of Treatment	-0.01378(0.04464)	0.7575	(0.11, 0.441)	-0.06358(0.03462)	0.0663	(0.234, 0.643)
Disease Stage						
Stage 1	-1.51198(0.35338)	<.0001	(0.28, 0.855)	-0.94594(0.25763)	0.0002	(0.358, 0.896)
Stage 2	-0.71431(0.28432)	0.012	(0.305, 0.861)	-0.56835(0.23423)	0.0152	(0.393, 0.935)
Stage 3	-0.66801(0.26476)	0.0116	(0.662, 3.531)	-0.50029(0.22078)	0.0235	(0.676, 2.534)
Creatinine Clearance						
<20 ml/min	0.42424(0.42721)	0.3207	(0.529, 2.935)	0.26925(0.33697)	0.4243	(0.687, 2.709)
20 - 40 ml/min	0.22(0.43713)	0.6148	(0.329, 1.575)	0.31072(0.35001)	0.3747	(0.476, 1.59)
40 - <60 ml/min	-0.32868(0.39932)	0.4105	(0.27, 1.201)	-0.13921(0.30759)	0.6509	(0.461, 1.448)
60 - <80 ml/min	-0.56252(0.38053)	0.1393	(0.395, 0.91)	-0.20243(0.29202)	0.4882	(0.451, 0.803)
Beta 2 Microglobulines						
<3.5 mg/dl	-0.51097(0.21263)	0.0163	(0.408, 1.367)	-0.50831(0.14717)	0.0006	(0.475, 1.113)
3.5 - 5 mg/dl	-0.29152(0.30834)	0.3444	(0.739, 1.552)	-0.31845(0.21724)	0.1427	(0.662, 1.211)
>5 mg/dl	0.06828(0.18948)	0.7186	(0.904, 1.076)	-0.11016(0.15393)	0.4742	(0.877, 1.004)

Table 9: Estimates (se) for Cox PH model adjusting for stratified propensity score

Covariate	OS			PFS		
	Estimate (se)	P-Value	95% Hazard Ratio Confidence Limits	Estimate (se)	P-Value	95% Hazard Ratio Confidence Limits
Baseline Treatment (Combination)	0.31069(0.17357)	0.0734	(0.971, 1.917)	-0.07509(0.1236)	0.5435	(0.728, 1.182)
PS Stratum	-0.01396(0.05351)	0.7943	(0.888, 1.095)	-0.03965(0.04015)	0.3234	(0.888, 1.04)

Table 10: Estimates (Empirical se) for weighted Cox PH with IPTW on propensity scores for treatment at baseline

Covariate	OS			PFS		
	Estimate (se)	P-Value	95% Confidence Limits	Estimate (se)	P-Value	95% Confidence Limits
Intercept	-1.7992(0.4872)	0.0002	(-2.754, -0.8443)	-0.5595(0.4172)	0.1799	(-1.3772, 0.2582)
Baseline Treatment(Combination)	0.1239(0.19)	0.5143	(-0.2485, 0.4964)	-0.1202(0.1431)	0.4011	(-0.4007, 0.1603)
Time Interval 2	0.6733(0.327)	0.0395	(0.0323, 1.3143)	1.2362(0.1848)	<.0001	(0.8741, 1.5983)
Time Interval 3	0.8913(0.2652)	0.0008	(0.3715, 1.4111)	1.3851(0.2183)	<.0001	(0.9572, 1.813)
Time Interval 4	1.0117(0.3208)	0.0016	(0.383, 1.6404)	0.4371(0.3138)	0.1636	(-0.1779, 1.052)
Time Interval 5	0.9029(0.4186)	0.031	(0.0826, 1.7233)	0.9817(0.3775)	0.0093	(0.2417, 1.7216)
Time Interval 6	0.6379(0.3677)	0.0828	(-0.0827, 1.3585)	0.2087(0.4521)	0.6443	(-0.6774, 1.0948)
Line of Treatment	-0.0045(0.0536)	0.9327	(-0.1097, 0.1006)	-0.0518(0.0446)	0.2452	(-0.1391, 0.0355)
Albumine						
>3.5 mg/dl	0.2705(0.3215)	0.4001	(-0.3596, 0.9007)	0.137(0.2216)	0.5362	(-0.2972, 0.5713)
<=3.5 mg/dl	0.8501(0.3294)	0.0098	(0.2046, 1.4956)	0.3548(0.2362)	0.1331	(-0.1082, 0.8178)
Gender (Female)	-0.1252(0.1802)	0.4873	(-0.4784, 0.2281)	-0.1086(0.1497)	0.468	(-0.402, 0.1848)
Disease Stage						
Stage 1	-1.8209(0.411)	<.0001	(-2.6265, -1.0154)	-1.133(0.3447)	0.001	(-1.8085, -0.4574)
Stage 2	-0.5509(0.3145)	0.0798	(-1.1674, 0.0655)	-0.7067(0.3094)	0.0224	(-1.313, -0.1003)
Stage 3	-0.6584(0.3041)	0.0304	(-1.2545, -0.0623)	-0.532(0.2932)	0.0696	(-1.1066, 0.0426)
Creatinine Clearance						
<20 ml/min	0.1053(0.6601)	0.8733	(-1.1886, 1.3992)	-0.0166(0.3924)	0.9662	(-0.7858, 0.7525)
20 - 40 ml/min	0.1535(0.5151)	0.7657	(-0.8562, 1.1632)	-0.0172(0.3667)	0.9625	(-0.7359, 0.7014)
40 - <60 ml/min	-0.6728(0.4999)	0.1783	(-1.6527, 0.307)	-0.3137(0.3163)	0.3212	(-0.9337, 0.3062)
60 - <80 ml/min	-1.0294(0.4533)	0.0232	(-1.9178, -0.1409)	-0.5447(0.2818)	0.0532	(-1.0969, 0.0076)
Beta 2 Microglobulines						
<3.5 mg/dl	-0.2057(0.2978)	0.4897	(-0.7895, 0.378)	-0.4984(0.2307)	0.0307	(-0.9505, -0.0463)
3.5 - 5 mg/dl	-0.2237(0.3202)	0.4849	(-0.8513, 0.404)	-0.4215(0.2378)	0.0763	(-0.8877, 0.0446)
>5 mg/dl	-0.0219(0.2389)	0.9268	(-0.4902, 0.4463)	-0.1874(0.1996)	0.348	(-0.5786, 0.2039)

Table 11: Estimates (Empirical se) for un-weighted logistic regression model with baseline treatment for panel data

Covariate	*OS*			*PFS*		
	Estimate (se)	*P-Value*	*95% Confidence Limits*	*Estimate (se)*	*P-Value*	*95% Confidence Limits*
Intercept	-1.8282(0.4682)	<.0001	(-2.7458,-0.9106)	-0.5166(0.397)	0.1932	(-1.2947, 0.2616)
Baseline Treatment(Combination)	0.2486(0.1858)	0.1809	(-0.1156, 0.6129)	-0.1308(0.1377)	0.3425	(-0.4007, 0.1392)
Time Interval 2	0.705(0.2051)	0.0006	(0.3029, 1.1071)	1.1741(0.1492)	<.0001	(0.8816, 1.4665)
Time Interval 3	0.8897(0.2126)	<.0001	(0.473, 1.3064)	1.2618(0.1759)	<.0001	(0.917, 1.6065)
Time Interval 4	0.8468(0.2424)	0.0005	(0.3716, 1.3219)	0.652(0.2436)	0.0074	(0.1747, 1.1294)
Time Interval 5	0.6613(0.3034)	0.0293	(0.0667, 1.2559)	0.9756(0.2978)	0.0011	(0.3921, 1.5592)
Time Interval 6	0.649(0.332)	0.0506	(-0.0018, 1.2998)	0.1905(0.4188)	0.6492	(-0.6303, 1.0113)
Albumine						
>3.5 mg/dl	0.0185(0.2581)	0.943	(-0.4874, 0.5244)	0.0007(0.1916)	0.997	(-0.3749, 0.3763)
<=3.5 mg/dl	0.5613(0.262)	0.0322	(0.0478, 1.0749)	0.3349(0.2063)	0.1046	(-0.0696, 0.7393)
Gender (Female)	-0.031(0.1571)	0.8436	(-0.3389, 0.2769)	-0.0843(0.124)	0.4963	(-0.3274, 0.1587)
Line of treatments	-0.0193(0.0463)	0.6764	(-0.11, 0.0714)	-0.0843(0.0387)	0.0294	(-0.1601, -0.0085)
Disease Stage						
Stage 1	-1.6204(0.3698)	<.0001	(-2.3451, -0.8957)	-1.0834(0.322)	0.0008	(-1.7145, -0.4523)
Stage 2	-0.7666(0.2899)	0.0082	(-1.3348, -0.1985)	-0.6881(0.2933)	0.019	(-1.2629,-0.1133)
Stage 3	-0.6811(0.2703)	0.0117	(-1.2108, -0.1514)	-0.5699(0.2772)	0.0398	(-1.1133, -0.0265)
Creatinine Clearance						
<20 ml/min	0.4116(0.4544)	0.3651	(-0.4791, 1.3023)	0.1849(0.3555)	0.603	(-0.5119, 0.8817)
20 - 40 ml/min	0.2265(0.4479)	0.613	(-0.6513, 1.1043)	0.2403(0.3467)	0.4882	(-0.4392, 0.9199)
40 - <60 ml/min	-0.392(0.3936)	0.3193	(-1.1634, 0.3794)	-0.3053(0.3066)	0.3193	(-0.9063, 0.2956)
60 - <80 ml/min	-0.6225(0.3642)	0.0874	(-1.3363, 0.0914)	-0.3622(0.2837)	0.2018	(-0.9182, 0.1939)
Beta 2 Microglobulines						
<3.5 mg/dl	-0.5035(0.223)	0.024	(-0.9406, -0.0663)	-0.5334(0.1645)	0.0012	(-0.8559, -0.2109)
3.5 - 5 mg/dl	-0.3119(0.3091)	0.3131	(-0.9178, 0.294)	-0.3673(0.2247)	0.1021	(-0.8077, 0.0731)
>5 mg/dl	0.0859(0.209)	0.681	(-0.3237, 0.4956)	-0.0725(0.1758)	0.6801	(-0.4171, 0.2721)

Table 12: Estimates (Empirical se) for un-weighted logistic regression with treatment history for panel data

Covariate	*OS*			*PFS*		
	Estimate (se)	*P-Value*	*95% Confidence Limits*	*Estimate (se)*	*P-Value*	*95% Confidence Limits*
Intercept	-1.4361(0.4672)	0.0021	(-2.3519, -0.5204)	-0.5074(0.3904)	0.1937	(-1.2727, 0.2578)
Treatment History (Combination)	-0.3843(0.1614)	0.0173	(-0.7007, -0.0678)	-0.1065(0.1345)	0.4286	(-0.3701, 0.1572)
Time Interval 2	0.6095(0.2055)	0.003	(0.2066, 1.0123)	1.1555(0.155)	<.0001	(0.8517, 1.4592)
Time Interval 3	0.7929(0.216)	0.0002	(0.3694, 1.2163)	1.2383(0.1808)	<.0001	(0.8839, 1.5927)
Time Interval 4	0.7444(0.2474)	0.0026	(0.2595, 1.2293)	0.6203(0.2452)	0.0114	(0.1398, 1.1008)
Time Interval 5	0.5554(0.3089)	0.0722	(-0.0501, 1.1608)	0.9536(0.2991)	0.0014	(0.3675, 1.5398)
Time Interval 6	0.5208(0.33)	0.1145	(-0.1259, 1.1675)	0.1839(0.4222)	0.6631	(-0.6435, 1.0113)
Albumine						
>3.5 mg/dl	0.0359(0.2564)	0.8886	(-0.4666, 0.5384)	-0.0874(0.039)	0.025	(-0.1639, -0.011)
<=3.5 mg/dl	0.5547(0.2598)	0.0327	(0.0455, 1.0639)	0.0151(0.1926)	0.9377	(-0.3624, 0.3925)
Gender (Female)	-0.0252(0.1574)	0.8729	(-0.3338, 0.2834)	0.3304(0.2068)	0.1102	(-0.075, 0.7357)
Line of Treatments	-0.0207(0.0474)	0.662	(-0.1136, 0.0722)	-0.0748(0.1237)	0.5456	(-0.3172, 0.1677)
Disease Stage						
Stage 1	-1.616(0.3748)	<.0001	(-2.3506, -0.8815)	-1.1088(0.3169)	0.0005	(-1.73, -0.4877)
Stage 2	-0.7698(0.2883)	0.0076	(-1.3348, -0.2047)	-0.728(0.2887)	0.0117	(-1.2939, -0.1621)
Stage 3	-0.6761(0.2673)	0.0114	(-1.2001, -0.1521)	-0.6117(0.2702)	0.0236	(-1.1412, -0.0822)
Creatinine Clearance						
<20 ml/min	0.4702(0.4608)	0.3075	(-0.4329, 1.3733)	0.1795(0.3533)	0.6113	(-0.5128, 0.8719)
20 - 40 ml/min	0.2554(0.4506)	0.5708	(-0.6277, 1.1386)	0.2145(0.3463)	0.5357	(-0.4643, 0.8933)
40 - <60 ml/min	-0.3803(0.3916)	0.3315	(-1.1478, 0.3873)	-0.3056(0.3076)	0.3204	(-0.9085, 0.2973)
60 - <80 ml/min	-0.6377(0.3629)	0.0789	(-1.349, 0.0737)	-0.373(0.2842)	0.1895	(-0.9301, 0.1842)
Beta 2 Microglobulines						
<3.5 mg/dl	-0.491(0.2175)	0.024	(-0.9172, -0.0648)	-0.5599(0.1625)	0.0006	(-0.8783, -0.2414)
3.5 - 5 mg/dl	-0.2797(0.305)	0.3591	(-0.8774, 0.318)	-0.3687(0.2245)	0.1005	(-0.8087, 0.0712)
>5 mg/dl	0.0795(0.2117)	0.7073	(-0.3355, 0.4945)	-0.0764(0.1753)	0.6629	(-0.4201, 0.2672)

Table 13: Estimates (Empirical se) for weighted Cox PH model for weights estimated under assumption 1

Covariate	OS			PFS		
	Estimate (se)	P-Value	95% Confidence Limits	Estimate (se)	P-Value	95% Confidence Limits
Intercept	-1.1111(0.5516)	0.044	(-2.1923, -0.0299)	-0.4646(0.4564)	0.3087	(-1.3592, 0.4299)
Treatment History (Combination)	-0.3908(0.1818)	0.0316	(-0.7471, -0.0345)	-0.3209(0.1494)	0.0317	(-0.6136, -0.0281)
Time Interval 2	0.672(0.2129)	0.0016	(0.2547, 1.0892)	1.1351(0.1588)	<.0001	(0.8238, 1.4463)
Time Interval 3	0.8535(0.2226)	0.0001	(0.4172, 1.2898)	1.2128(0.1854)	<.0001	(0.8495, 1.5762)
Time Interval 4	0.8025(0.2551)	0.0017	(0.3024, 1.3025)	0.5611(0.2506)	0.0251	(0.07, 1.0522)
Time Interval 5	0.5976(0.3158)	0.0584	(-0.0213, 1.2165)	0.8972(0.3077)	0.0036	(0.2941, 1.5003)
Time Interval 6	0.5701(0.3385)	0.0922	(-0.0935, 1.2336)	0.1371(0.4283)	0.7488	(-0.7023, 0.9765)
Albumine						
>3.5 mg/dl	0.1176(0.2806)	0.6751	(-0.4323, 0.6675)	0.1487(0.2036)	0.4651	(-0.2504, 0.5479)
<=3.5 mg/dl	0.7619(0.2845)	0.0074	(0.2044, 1.3195)	0.5643(0.218)	0.0096	(0.137, 0.9915)
Gender (Female)	-0.159(0.1808)	0.3793	(-0.5134, 0.1954)	-0.1219(0.1375)	0.3751	(-0.3914, 0.1475)
Line of Treatment	-0.0552(0.0513)	0.2814	(-0.1557, 0.0452)	-0.0672(0.0422)	0.1117	(-0.1499, 0.0156)
Disease Stage						
Stage 1	-1.6666(0.4117)	<.0001	(-2.4736, -0.8596)	-1.0541(0.3162)	0.0009	(-1.6739, -0.4343)
Stage 2	-1.0165(0.3419)	0.0029	(-1.6866, -0.3465)	-0.7604(0.2856)	0.0078	(-1.3202, -0.2006)
Stage 3	-0.8588(0.3109)	0.0057	(-1.4681, -0.2495)	-0.7146(0.2654)	0.0071	(-1.2348, -0.1944)
Creatinine Clearance						
<20 ml/min	0.7159(0.4964)	0.1492	(-0.257, 1.6889)	0.3228(0.4058)	0.4264	(-0.4726, 1.1182)
20 - 40 ml/min	0.3313(0.4968)	0.5049	(-0.6424, 1.3049)	0.224(0.3942)	0.57	(-0.5487, 0.9966)
40 - <60 ml/min	-0.2988(0.4535)	0.5099	(-1.1877, 0.59)	-0.2333(0.3578)	0.5143	(-0.9346, 0.4679)
60 - <80 ml/min	-0.7048(0.4219)	0.0948	(-1.5317, 0.1221)	-0.4502(0.3364)	0.1808	(-1.1095, 0.2091)
Beta 2 Microglobulines						
<3.5 mg/dl	-0.5954(0.2505)	0.0175	(-1.0865, -0.1044)	-0.5422(0.1807)	0.0027	(-0.8964, -0.1879)
3.5 - 5 mg/dl	-0.5688(0.3598)	0.1139	(-1.2739, 0.1364)	-0.5205(0.2446)	0.0334	(-1, -0.041)
>5 mg/dl	0.0301(0.2333)	0.8974	(-0.4271, 0.4873)	-0.0602(0.1856)	0.7457	(-0.424, 0.3036)

Table 14: Estimates (Empirical se) for weighted Cox PH model for weights estimated under assumption 2

Covariate	OS			PFS		
	Estimate (se)	P-Value	95% Confidence Limits	Estimate (se)	P-Value	95% Confidence Limits
Intercept	-1.1088(0.5516)	0.0444	(-2.19, -0.0277)	-0.4814(0.4559)	0.291	(-1.3748, 0.4121)
Treatment History (Combination)	-0.3895(0.1817)	0.0321	(-0.7456, -0.0334)	-0.2955(0.1492)	0.0477	(-0.588, -0.003)
Time Interval 2	0.6726(0.2128)	0.0016	(0.2555, 1.0897)	1.1445(0.1586)	<.0001	(0.8337, 1.4553)
Time Interval 3	0.8549(0.2226)	0.0001	(0.4186, 1.2912)	1.2233(0.1852)	<.0001	(0.8603, 1.5863)
Time Interval 4	0.8041(0.2551)	0.0016	(0.3041, 1.304)	0.5684(0.2502)	0.0231	(0.078, 1.0587)
Time Interval 5	0.599(0.3157)	0.0578	(-0.0197, 1.2178)	0.9042(0.3074)	0.0033	(0.3016, 1.5067)
Time Interval 6	0.5711(0.3385)	0.0915	(-0.0923, 1.2346)	0.1455(0.4277)	0.7337	(-0.6927, 0.9837)
Albumine						
>3.5 mg/dl	0.1157(0.2809)	0.6805	(-0.4349, 0.6662)	0.1445(0.204)	0.4789	(-0.2554, 0.5443)
<=3.5 mg/dl	0.7609(0.2849)	0.0076	(0.2026, 1.3192)	0.5562(0.2184)	0.0109	(0.1281, 0.9844)
Gender (Female)	-0.1602(0.1809)	0.3758	(-0.5147, 0.1943)	-0.1227(0.1374)	0.3717	(-0.392, 0.1466)
Line of Treatment	-0.0554(0.0513)	0.2801	(-0.1558, 0.0451)	-0.0679(0.0421)	0.1065	(-0.1504, 0.0146)
Disease Stage						
Stage 1	-1.6675(0.4117)	<.0001	(-2.4743, -0.8607)	-1.0557(0.3172)	0.0009	(-1.6773, -0.434)
Stage 2	-1.0157(0.3418)	0.003	(-1.6856, -0.3458)	-0.7604(0.2866)	0.008	(-1.3221, -0.1987)
Stage 3	-0.8581(0.3109)	0.0058	(-1.4674, -0.2488)	-0.7124(0.2665)	0.0075	(-1.2347, -0.19)
Creatinine Clearance						
<20 ml/min	0.7144(0.4969)	0.1505	(-0.2594, 1.6882)	0.3252(0.405)	0.422	(-0.4686, 1.119)
20 - 40 ml/min	0.3299(0.4971)	0.507	(-0.6445, 1.3043)	0.228(0.3937)	0.5624	(-0.5435, 0.9996)
40 - <60 ml/min	-0.3002(0.454)	0.5085	(-1.1901, 0.5897)	-0.23(0.357)	0.5194	(-0.9298, 0.4697)
60 - <80 ml/min	-0.7065(0.4223)	0.0943	(-1.5343, 0.1212)	-0.4427(0.3359)	0.1875	(-1.1009, 0.2156)
Beta 2 Microglobulines						
<3.5 mg/dl	-0.5935(0.2507)	0.0179	(-1.0849, -0.1022)	-0.5339(0.1813)	0.0032	(-0.8892, -0.1787)
3.5 - 5 mg/dl	-0.5689(0.3594)	0.1135	(-1.2734, 0.1356)	-0.5207(0.2443)	0.0331	(-0.9996, -0.0418)
>5 mg/dl	0.0312(0.2333)	0.8937	(-0.4261, 0.4884)	-0.0593(0.1854)	0.7493	(-0.4227, 0.3041)

Table 15: Estimates (Empirical se) for weighted Cox PH model for weights estimated under assumption 3

Covariate	OS			PFS		
	Estimate (se)	P-Value	95% Confidence Limits	Estimate (se)	P-Value	95% Confidence Limits
Intercept	-1.2373(0.7558)	0.1016	(-2.7186, 0.2441)	-0.4466(0.5814)	0.4424	(-1.5861, 0.6929)
Treatment History (Combination)	-0.4111(0.1754)	0.0191	(-0.755, -0.0673)	-0.2358(0.1489)	0.1133	(-0.5277, 0.0561)
Time Interval 2	0.6115(0.2256)	0.0067	(0.1693, 1.0538)	1.3926(0.1768)	<.0001	(1.046, 1.7391)
Time Interval 3	0.8751(0.2332)	0.0002	(0.418, 1.3322)	1.4047(0.1993)	<.0001	(1.0141, 1.7954)
Time Interval 4	0.8342(0.2649)	0.0016	(0.3151, 1.3533)	0.6582(0.262)	0.012	(0.1446, 1.1718)
Time Interval 5	0.6573(0.321)	0.0406	(0.0281, 1.2865)	1.1136(0.3188)	0.0005	(0.4887, 1.7385)
Time Interval 6	0.5473(0.3412)	0.1087	(-0.1215, 1.2161)	0.3204(0.4419)	0.4684	(-0.5457, 1.1865)
Albumine						
>3.5 mg/dl	0.2382(0.3186)	0.4546	(-0.3862, 0.8626)	0.4335(0.2458)	0.0778	(-0.0483, 0.9154)
<=3.5 mg/dl	0.8517(0.3247)	0.0087	(0.2154, 1.488)	0.786(0.2598)	0.0025	(0.2767, 1.2953)
Gender (Female)	-0.134(0.2235)	0.5487	(-0.5721, 0.304)	-0.2022(0.1734)	0.2435	(-0.542, 0.1376)
Line of Treatment	-0.0225(0.058)	0.698	(-0.1363, 0.0912)	-0.0558(0.0434)	0.1984	(-0.1408, 0.0292)
Disease Stage						
Stage 1	-1.3209(0.4488)	0.0032	(-2.2006, -0.4413)	-1.0897(0.3751)	0.0037	(-1.8249, -0.3545)
Stage 2	-0.988(0.4152)	0.0173	(-1.8018, -0.1742)	-0.7261(0.3557)	0.0412	(-1.4232, -0.0291)
Stage 3	-0.8424(0.3605)	0.0195	(-1.5491, -0.1358)	-0.9051(0.3291)	0.0059	(-1.55, -0.2602)
Creatinine Clearance						
<20 ml/min	0.4904(0.6238)	0.4318	(-0.7322, 1.713)	0.042(0.4968)	0.9327	(-0.9318, 1.0157)
20 - 40 ml/min	0.2828(0.6176)	0.6471	(-0.9278, 1.4933)	-0.1171(0.4984)	0.8143	(-1.094, 0.8598)
40 - <60 ml/min	-0.4207(0.575)	0.4644	(-1.5476, 0.7063)	-0.4316(0.4597)	0.3478	(-1.3325, 0.4694)
60 - <80 ml/min	-0.8685(0.547)	0.1123	(-1.9406, 0.2035)	-0.6029(0.4393)	0.1699	(-1.4639, 0.258)
Beta 2 Microglobulines						
<3.5 mg/dl	-0.8147(0.3191)	0.0107	(-1.4401, -0.1894)	-0.7635(0.2214)	0.0006	(-1.1975, -0.3295)
3.5 - 5 mg/dl	-0.4855(0.4704)	0.302	(-1.4076, 0.4365)	-0.7265(0.306)	0.0176	(-1.3262, -0.1267)
>5 mg/dl	0.1291(0.2681)	0.6302	(-0.3963, 0.6545)	-0.0175(0.2259)	0.9384	(-0.4602, 0.4253)

Table 16: Estimates (Empirical se) for weighted Cox PH model for weights estimated under assumption 4

Covariate	OS			PFS		
	Estimate (se)	P-Value	95% Confidence Limits	Estimate (se)	P-Value	95% Confidence Limits
Intercept	-1.1264(0.5772)	0.051	(-2.2577, 0.005)	-0.327(0.4819)	0.4974	(-1.2715, 0.6176)
Treatment History (Combination)	-0.4458(0.1885)	0.018	(-0.8152, -0.0764)	-0.5398(0.1532)	0.0004	(-0.8401, -0.2394)
Time Interval 2	0.6136(0.2183)	0.0049	(0.1858, 1.0414)	1.1018(0.1631)	<.0001	(0.782, 1.4215)
Time Interval 3	0.8266(0.2271)	0.0003	(0.3814, 1.2718)	1.165(0.1888)	<.0001	(0.795, 1.5351)
Time Interval 4	0.77(0.2591)	0.003	(0.2621, 1.2779)	0.4913(0.2554)	0.0544	(-0.0094, 0.992)
Time Interval 5	0.5698(0.3187)	0.0738	(-0.0548, 1.1944)	0.8711(0.3138)	0.0055	(0.256, 1.4862)
Time Interval 6	0.5165(0.3425)	0.1316	(-0.1549, 1.1878)	0.0688(0.4346)	0.8741	(-0.783, 0.9207)
Albumine						
>3.5 mg/dl	0.2145(0.273)	0.432	(-0.3206, 0.7496)	0.2238(0.2095)	0.2854	(-0.1868, 0.6345)
<=3.5 mg/dl	0.8423(0.2754)	0.0022	(0.3025, 1.382)	0.6707(0.2241)	0.0028	(0.2314, 1.11)
Gender (Female)	-0.1427(0.1853)	0.4413	(-0.5059, 0.2205)	-0.1155(0.1432)	0.4202	(-0.3962, 0.1653)
Line of Treatment	-0.0509(0.0521)	0.3289	(-0.153, 0.0512)	-0.0611(0.0438)	0.1625	(-0.1469, 0.0247)
Disease Stage						
Stage 1	-1.6159(0.4212)	0.0001	(-2.4415, -0.7903)	-1.012(0.3259)	0.0019	(-1.6508, -0.3733)
Stage 2	-1.0295(0.3512)	0.0034	(-1.7178, -0.3411)	-0.7432(0.2963)	0.0121	(-1.3239, -0.1625)
Stage 3	-0.8979(0.3169)	0.0046	(-1.5189, -0.2769)	-0.7314(0.275)	0.0078	(-1.2704, -0.1924)
Creatinine Clearance						
<20 ml/min	0.7015(0.5012)	0.1616	(-0.2809, 1.6838)	0.2688(0.428)	0.53	(-0.5701, 1.1077)
20 - 40 ml/min	0.3361(0.5003)	0.5017	(-0.6445, 1.3167)	0.1461(0.416)	0.7254	(-0.6693, 0.9616)
40 - <60 ml/min	-0.2969(0.4547)	0.5137	(-1.1881, 0.5942)	-0.2691(0.3794)	0.4781	(-1.0127, 0.4744)
60 - <80 ml/min	-0.6998(0.4245)	0.0993	(-1.5319, 0.1322)	-0.5381(0.3563)	0.1309	(-1.2364, 0.1602)
Beta 2 Microglobulines						
<3.5 mg/dl	-0.6644(0.2571)	0.0098	(-1.1683, -0.1604)	-0.621(0.1845)	0.0008	(-0.9826, -0.2593)
3.5 - 5 mg/dl	-0.5904(0.3802)	0.1204	(-1.3356, 0.1547)	-0.5678(0.2556)	0.0263	(-1.0688, -0.0668)
>5 mg/dl	-0.0023(0.2373)	0.9923	(-0.4674,0.4628)	-0.091(0.1931)	0.6374	(-0.4695, 0.2875)

Table 17: Estimates (se) for Cox PH model for Censoring method at time of crossover

Covariate	*OS*			*PFS*		
	Estimate (se)	*P-Value*	*95% Hazard Ratio Confidence Limits*	*Estimate (se)*	*P-Value*	*95% Hazard Ratio Confidence Limits*
Baseline Treatment(Combination)	-0.16101(0.32751)	0.623	(0.448, 1.618)	-0.05832(0.21182)	0.7831	(0.623, 1.429)
Albumine						
>3.5 mg/dl	-0.5613(0.50002)	0.2616	(0.214, 1.52)	-0.06361(0.33707)	0.8503	(0.485, 1.817)
< =3.5 mg/dl	0.39619(0.46529)	0.3945	(0.597, 3.699)	0.49923(0.33294)	0.1338	(0.858, 3.164)
Gender (Female)	-0.26619(0.29603)	0.3686	(0.429, 1.369)	-0.25216(0.1958)	0.1978	(0.529, 1.141)
Line of Treatment	-0.01898(0.08177)	0.8165	(0.836, 1.152)	-0.05094(0.05628)	0.3654	(0.851, 1.061)
Disease Stage						
Stage 1	-1.25041(0.57686)	0.0302	(0.092, 0.887)	-0.78587(0.38909)	0.0434	(0.213, 0.977)
Stage 2	-1.14927(0.49386)	0.02	(0.12, 0.834)	-0.73404(0.36118)	0.0421	(0.236, 0.974)
Stage 3	-0.94749(0.44309)	0.0325	(0.163, 0.924)	-0.79511(0.33943)	0.0192	(0.232, 0.878)
Creatinine Clearance						
<20 ml/min	0.70756(0.6821)	0.2996	(0.533, 7.725)	0.28851(0.51547)	0.5757	(0.486, 3.665)
20 - 40 ml/min	0.63335(0.72749)	0.384	(0.453, 7.84)	0.37305(0.53917)	0.489	(0.505, 4.178)
40 - <60 ml/min	-0.29643(0.67722)	0.6616	(0.197, 2.804)	-0.24185(0.48522)	0.6182	(0.303, 2.032)
60 - <80 ml/min	-0.97008(0.65489)	0.1385	(0.105, 1.368)	-0.6467(0.46558)	0.1648	(0.21, 1.304)
Beta 2 Microglobulines						
<3.5 mg/dl	-0.59505(0.51361)	0.2466	(0.202, 1.509)	-0.30511(0.25896)	0.2387	(0.444, 1.224)
3.5 - 5 mg/dl	0.19049(0.55101)	0.7296	(0.411, 3.562)	-0.53027(0.43491)	0.2227	(0.251, 1.38)
>5 mg/dl	0.11034(0.34334)	0.7479	(0.57, 2.189)	-0.02743(0.24337)	0.9103	(0.604, 1.568)

Table 18: Estimates (Empirical se) for weighted Cox PH model with non-stabilized weights(IPTW)

Covariate	OS			PFS		
	Estimate (se)	P-Value	95% Confidence Limits	Estimate (se)	P-Value	95% Confidence Limits
Intercept	-1.0027(0.6178)	0.1046	(-2.2135, 0.2081)	0.1659(0.5634)	0.7685	(-0.9384, 1.2701)
Treatment History (Combination)	-0.463(0.2413)	0.055	(-0.9358, 0.0099)	-0.1468(0.1711)	0.391	(-0.4822, 0.1886)
Time Interval 2	-0.1288(0.3192)	0.6866	(-0.7544, 0.4968)	0.6799(0.2065)	0.001	(0.2752, 1.0846)
Time Interval 3	0.2939(0.3418)	0.3899	(-0.376, 0.9637)	0.6504(0.2303)	0.0047	(0.1989, 1.1018)
Time Interval 4	-0.0254(0.406)	0.9501	(-0.8212, 0.7704)	0.4211(0.3142)	0.1801	(-0.1947, 1.0369)
Time Interval 5	0.0967(0.5049)	0.8481	(-0.8928, 1.0863)	0.5194(0.3668)	0.1568	(-0.1995, 1.2382)
Time Interval 6	-0.6477(0.4594)	0.1586	(-1.5481, 0.2527)	-0.2956(0.5392)	0.5836	(-1.3523, 0.7612)
Albumine						
>3.5 mg/dl	-0.6252(0.4279)	0.144	(-1.4639, 0.2136)	-0.6072(0.2649)	0.0219	(-1.1264, -0.088)
<=3.5 mg/dl	-0.2785(0.4453)	0.5317	(-1.1513, 0.5943)	-0.3069(0.2824)	0.2772	(-0.8604, 0.2466)
Gender (Female)	-0.1675(0.2727)	0.539	(-0.7021, 0.367)	-0.0818(0.1826)	0.6542	(-0.4397, 0.2761)
Line of Treatment	-0.1658(0.0783)	0.0342	(-0.3192, -0.0124)	-0.1713(0.056)	0.0022	(-0.281, -0.0616)
Disease Stage						
Stage 1	-0.8559(0.537)	0.111	(-1.9084, 0.1966)	-0.7059(0.5234)	0.1775	(-1.7317, 0.32)
Stage 2	-0.3776(0.409)	0.3558	(-1.1792, 0.424)	-0.3799(0.4785)	0.4272	(-1.3178, 0.5579)
Stage 3	-0.2817(0.3435)	0.4121	(-0.955, 0.3915)	-0.4365(0.4674)	0.3503	(-1.3526, 0.4795)
Creatinine Clearance						
<20 ml/min	1.0667(0.6236)	0.0871	(-0.1555, 2.2888)	0.0323(0.4653)	0.9447	(-0.8797, 0.9443)
20 - 40 ml/min	1.2656(0.7037)	0.0721	(-0.1138, 2.6449)	0.6575(0.463)	0.1556	(-0.2499, 1.5649)
40 - <60 ml/min	0.9456(0.5425)	0.0813	(-0.1176, 2.0089)	-0.0415(0.4398)	0.9249	(-0.9035, 0.8206)
60 - <80 ml/min	0.4051(0.4892)	0.4077	(-0.5538, 1.364)	-0.1211(0.4101)	0.7677	(-0.925, 0.6827)
Beta 2 Microglobulines						
<3.5 mg/dl	-0.3059(0.3324)	0.3573	(-0.9573, 0.3455)	-0.5602(0.2284)	0.0142	(-1.0078, -0.1126)
3.5 - 5 mg/dl	-0.2769(0.3733)	0.4582	(-1.0086, 0.4548)	-0.2138(0.2994)	0.4752	(-0.8006, 0.373)
>5 mg/dl	-0.0514(0.341)	0.8801	(-0.7197, 0.6168)	-0.2432(0.2532)	0.3369	(-0.7395, 0.2532)

Table 19: Estimates (Empirical se) for weighted Cox PH model with non-stabilized weights (IPCW)

Covariate	OS			PFS		
	Estimate (se)	P-Value	95% Confidence Limits	Estimate (se)	P-Value	95% Confidence Limits
Intercept	-1.1506(0.7006)	0.1005	(-2.5237,0.2225)	-1.2698(0.5961)	0.0331	(-2.4381, -0.1015)
Treatment History (Combination)	-0.6971(0.2755)	0.0114	(-1.237, -0.1572)	0.3774(0.2606)	0.1476	(-0.1335, 0.8882)
Time Interval 2	0.9882(0.2635)	0.0002	(0.4718, 1.5047)	1.4002(0.186)	<.0001	(1.0357, 1.7647)
Time Interval 3	1.0984(0.2844)	0.0001	(0.541, 1.6558)	1.5687(0.2188)	<.0001	(1.1397, 1.9976)
Time Interval 4	1.0112(0.2979)	0.0007	(0.4273, 1.5952)	0.8141(0.2828)	0.004	(0.2599, 1.3683)
Time Interval 5	0.7156(0.3739)	0.0557	(-0.0173, 1.4484)	1.7128(0.424)	<.0001	(0.8818, 2.5438)
Time Interval 6	0.5776(0.3617)	0.1103	(-0.1314, 1.2866)	1.0753(0.5008)	0.0318	(0.0937, 2.0569)
Line of treatments	0.1241(0.0685)	0.07	(-0.0101, 0.2583)	-0.0633(0.0518)	0.2211	(-0.1648, 0.0381)
Albumine						
>3.5 mg/dl	0.4791(0.3731)	0.1991	(-0.2522, 1.2105)	0.0554(0.2561)	0.8287	(-0.4465,0.5573)
<=3.5 mg/dl	1.0061(0.4075)	0.0136	(0.2074, 1.8049)	0.3767(0.2874)	0.19	(-0.1867, 0.94)
Gender (Female)	-0.1708(0.241)	0.4786	(-0.6432, 0.3016)	-0.0991(0.2434)	0.6839	(-0.5762, 0.378)
Disease Stage						
Stage 1	-2.9096(0.5465)	<.0001	(-3.9807, -1.8385)	-1.4041(0.5367)	0.0089	(-2.4561, -0.3521)
Stage 2	-0.5129(0.3912)	0.1898	(-1.2796, 0.2538)	-0.458(0.4654)	0.3251	(-1.3702, 0.4543)
Stage 3	-0.7272(0.3604)	0.0436	(-1.4336, -0.0209)	-0.1133(0.4305)	0.7924	(-0.9571, 0.7305)
Creatinine Clearance						
<20 ml/min	-0.3675(0.7391)	0.6191	(-1.8162, 1.0812)	0.4504(0.3865)	0.2439	(-0.3071, 1.208)
20 - 40 ml/min	-0.9429(0.6863)	0.1695	(-2.2881, 0.4023)	0.2686(0.4127)	0.5151	(-0.5402, 1.0775)
40 - <60 ml/min	-1.3763(0.6182)	0.026	(-2.5881, -0.1646)	-0.2952(0.3762)	0.4327	(-1.0325, 0.4422)
60 - <80 ml/min	-1.437(0.5759)	0.0126	(-2.5656, -0.3083)	-0.2546(0.3267)	0.4358	(-0.8949, 0.3857)
Beta 2 Microglobulines						
<3.5 mg/dl	-0.5083(0.3195)	0.1116	(-1.1345, 0.1179)	-0.0979(0.2408)	0.6842	(-0.5699, 0.3741)
3.5 - 5 mg/dl	-0.2381(0.4183)	0.5692	(-1.058, 0.5817)	-0.5922(0.3811)	0.1202	(-1.3391, 0.1547)
>5 mg/dl	0.3809(0.3458)	0.2707	(-0.2969, 1.0588)	-0.1431(0.2858)	0.6167	(-0.7032, 0.4171)

SAS CODES

OS End Point

```
/*Fitting Ordinary Cox PH model
for OS Treat only*/
proc phreg data=Data_original;
model
time_to_death*status_death(0)=
A/risklimits;
run; quit;
/*Fitting Ordinary Cox PH model
Treat and covariates*/
proc phreg data=Data_original;
model
time_to_death*status_death(0)= A
Albumin_cat1 Albumin_cat2  Gender
MMStage_common1
MMStage_common2 MMStage_common3
creatclear_cat1 creatclear_cat2
creatclear_cat3 creatclear_cat4
B2Microglob_cat1
B2Microglob_cat2 B2Microglob_cat3
LOT/risklimits;

Albumin_cat1=(Albumin_cat=1);
Albumin_cat2=(Albumin_cat=2);
MMStage_common1=(MMStage_common=1)
;
MMStage_common2=(MMStage_common=2)
;
MMStage_common3=(MMStage_common=3)
;
creatclear_cat1=(creatclear_cat=1)
;
creatclear_cat2=(creatclear_cat=2)
;
creatclear_cat3=(creatclear_cat=3)
;
creatclear_cat4=(creatclear_cat=4)
;
B2Microglob_cat1=(B2Microglob_cat=
1);
B2Microglob_cat2=(B2Microglob_cat=
2);
B2Microglob_cat3=(B2Microglob_cat=
3);



Albumin_cat: test Albumin_cat1,
Albumin_cat2;
MMStage_common: test
MMStage_common1, MMStage_common2,
MMStage_common3;
creatclear_cat: test
creatclear_cat1, creatclear_cat2,
creatclear_cat3, creatclear_cat4;
B2Microglob_cat: test
B2Microglob_cat1,
B2Microglob_cat2,
B2Microglob_cat3;
run; quit;
```

```
/*Cox Model for adjusting for the
stratified propensity score*/
proc phreg data=PS_strata;
model
time_to_death*status_death(0)= A
stratum/risklimits;
run; quit;
/*Weighted Cox PH with IPTW based
on Propensity scores*/
proc genmod data=PSdata1
descending;
class PatientID Albumin_cat
MMStage_common creatclear_cat
B2Microglob_cat
Treat_A_BL(ref=first)/param=ref;
model Event = Treat_A_BL Time2
Time3 Time4 Time5 Time6 LOT
Albumin_cat Gender MMStage_common
creatclear_cat B2Microglob_cat/
link=logit dist=bin;
scwgt PS_weight;
repeated subject=PatientID/
type=Ind;
run;quit;

/*Un-Weighted logistic regression
with baseline treatment*/
proc genmod data=OS descending;
class PatientID Albumin_cat
MMStage_common creatclear_cat
B2Microglob_cat
Treat_A_BL(ref=first)/param=ref;
model Event = Treat_A_BL Time2
Time3 Time4 Time5 Time6 LOT
Albumin_cat Gender MMStage_common
creatclear_cat B2Microglob_cat/
link=logit dist=bin;
repeated subject=PatientID/
type=Ind;
run;quit;

/*Un-Weighted logistic regression
with treatment history*/
proc genmod data=OS descending;
class PatientID Albumin_cat
MMStage_common creatclear_cat
B2Microglob_cat;
model Event = Treat_A2 Time2 Time3
Time4 Time5 Time6 LOT Albumin_cat
Gender MMStage_common
creatclear_cat B2Microglob_cat/
link=logit dist=bin;
repeated subject=PatientID/
type=Ind;
run;quit;

/*Weighted Cox PH with non-
stabilized IPTW*/
proc genmod data=OS_Weight_IPTW
descending;
```

```
class PatientID Albumin_cat
MMStage_common creatclear_cat
B2Microglob_cat
Treat_A2(ref=first)/param=ref;
model Event = Treat_A2 Time2 Time3
Time4 Time5 Time6 LOT Albumin_cat
Gender MMStage_common
creatclear_cat B2Microglob_cat/
link=logit dist=bin;
scwgt nstb_weight;
repeated subject=PatientID/
type=Ind;
run;quit;

/*Weighted Cox PH with non-
stabilized IPCW*/
proc genmod data=OS_Weight_IPCW
descending;
class PatientID Albumin_cat
MMStage_common creatclear_cat
B2Microglob_cat
Treat_A2(ref=first)/param=ref;
model Event = Treat_A2 Time2 Time3
Time4 Time5 Time6 LOT Albumin_cat
Gender MMStage_common
creatclear_cat B2Microglob_cat/
link=logit dist=bin;
scwgt non_stb_weight;
repeated subject=PatientID/
type=Ind;
run;quit;

/*Weighted Cox PH with Weight 1
(Crossover probabilities)*/
proc genmod data=OS_Weight
descending;
class PatientID Albumin_cat
MMStage_common creatclear_cat
B2Microglob_cat
Treat_A2(ref=first)/param=ref;
model Event = Treat_A2 Time2 Time3
Time4 Time5 Time6 LOT Albumin_cat
Gender MMStage_common
creatclear_cat B2Microglob_cat/
link=logit dist=bin;
scwgt w1;
repeated subject=PatientID/
type=Ind;
run;quit;

/*Weighted Cox PH with Weight 2
(Crossover probabilities)*/
proc genmod data=OS_Weight
descending;
class PatientID Albumin_cat
MMStage_common creatclear_cat
B2Microglob_cat
Treat_A2(ref=first)/param=ref;
model Event = Treat_A2 Time2 Time3
Time4 Time5 Time6 LOT Albumin_cat
Gender MMStage_common
creatclear_cat B2Microglob_cat/
link=logit dist=bin;
scwgt w2;
repeated subject=PatientID/
type=Ind;
run;quit;

/*Weighted Cox PH with Weight 3
(Crossover probabilities)*/
proc genmod data=OS_Weight
descending;
class PatientID Albumin_cat
MMStage_common creatclear_cat
B2Microglob_cat
Treat_A2(ref=first)/param=ref;
model Event = Treat_A2 Time2 Time3
Time4 Time5 Time6 LOT Albumin_cat
Gender MMStage_common
creatclear_cat B2Microglob_cat/
link=logit dist=bin;
scwgt w3;
repeated subject=PatientID/
type=Ind;
run;quit;

/*Weighted Cox PH with Weight 4
(Crossover probabilities)*/
proc genmod data=OS_Weight
descending;
class PatientID Albumin_cat
MMStage_common creatclear_cat
B2Microglob_cat
Treat_A2(ref=first)/param=ref;
model Event = Treat_A2 Time2 Time3
Time4 Time5 Time6 LOT Albumin_cat
Gender MMStage_common
creatclear_cat B2Microglob_cat/
link=logit dist=bin;
scwgt w4;
repeated subject=PatientID/
type=Ind;
run;quit;
```

PFS End Point

```
/*Fitting Ordinary Cox PH model
for OS Treat only*/
proc phreg data=Data_original;
model time_to_PFS*status_PFS(0)=
A/risklimits;
run; quit;

/*Fitting Ordinary Cox PH model
with treat and covariate*/
proc phreg data=Data_original;
model time_to_PFS*status_PFS(0)= A
Albumin_cat1 Albumin_cat2 Gender
MMStage_common1
MMStage_common2 MMStage_common3
creatclear_cat1 creatclear_cat2
```

```
creatclear_cat3 creatclear_cat4
B2Microglob_cat1
B2Microglob_cat2 B2Microglob_cat3
LOT/risklimits;

Albumin_cat1=(Albumin_cat=1);
Albumin_cat2=(Albumin_cat=2);
MMStage_common1=(MMStage_common=1)
;
MMStage_common2=(MMStage_common=2)
;
MMStage_common3=(MMStage_common=3)
;
creatclear_cat1=(creatclear_cat=1)
;
creatclear_cat2=(creatclear_cat=2)
;
creatclear_cat3=(creatclear_cat=3)
;
creatclear_cat4=(creatclear_cat=4)
;
B2Microglob_cat1=(B2Microglob_cat=
1);
B2Microglob_cat2=(B2Microglob_cat=
2);
B2Microglob_cat3=(B2Microglob_cat=
3);
Albumin_cat: test Albumin_cat1,
Albumin_cat2;
MMStage_common: test
MMStage_common1, MMStage_common2,
MMStage_common3;
creatclear_cat: test
creatclear_cat1, creatclear_cat2,
creatclear_cat3, creatclear_cat4;
B2Microglob_cat: test
B2Microglob_cat1,
B2Microglob_cat2,
B2Microglob_cat3;
run; quit;

/*Cox Model for adjusting for the
stratified propensity score*/
proc phreg data=PS_strata;
model time_to_PFS*status_PFS(0)= A
stratum/risklimits;
run; quit;

/*Weighted Cox PH with IPTW based
on Propensity scores*/
proc genmod data=PSdata3
descending;
class PatientID Albumin_cat
MMStage_common creatclear_cat
B2Microglob_cat
Treat_A_BL(ref=first)/param=ref;
model Event = Treat_A_BL Time2
Time3 Time4 Time5 Time6 LOT
Albumin_cat Gender MMStage_common
creatclear_cat B2Microglob_cat/
link=logit dist=bin;
scwgt PS_weight;
repeated subject=PatientID/
type=Ind;
run;quit;

/*Un-Weighted logistic regression
with baseline treatment*/
proc genmod data=PFS descending;
class PatientID Albumin_cat
MMStage_common creatclear_cat
B2Microglob_cat
Treat_A_BL(ref=first)/param=ref;
model Event = Treat_A_BL Time2
Time3 Time4 Time5 Time6 LOT
Albumin_cat Gender MMStage_common
creatclear_cat B2Microglob_cat/
link=logit dist=bin;
repeated subject=PatientID/
type=Ind;
run;quit;

/*Un-Weighted logistic regression
with treatment history*/
proc genmod data=PFS descending;
class PatientID Albumin_cat
MMStage_common creatclear_cat
B2Microglob_cat
Treat_A2(ref=first)/param=ref;
model Event = Treat_A2 Time2 Time3
Time4 Time5 Time6 LOT Albumin_cat
Gender MMStage_common
creatclear_cat B2Microglob_cat/
link=logit dist=bin;
repeated subject=PatientID/
type=Ind;
run;quit;

/*Weighted Cox PH with non-
stabilized IPTW*/
proc genmod data=PFS_Weight_IPTW
descending;
class PatientID Albumin_cat
MMStage_common creatclear_cat
B2Microglob_cat
Treat_A2(ref=first)/param=ref;
model Event = Treat_A2 Time2 Time3
Time4 Time5 Time6 LOT Albumin_cat
Gender MMStage_common
creatclear_cat B2Microglob_cat/
link=logit dist=bin;
scwgt nstb_weight;
repeated subject=PatientID/
type=Ind;
run;quit;


/*Weighted Cox PH with non-
stabilized IPCW*/
proc genmod data=PFS_Weight_IPCW
descending;
```

```
class PatientID Albumin_cat
MMStage_common creatclear_cat
B2Microglob_cat
Treat_A2(ref=first)/param=ref;
model Event = Treat_A2 Time2 Time3
Time4 Time5 Time6 LOT Albumin_cat
Gender MMStage_common
creatclear_cat B2Microglob_cat/
link=logit dist=bin;
scwgt non_stb_weight;
repeated subject=PatientID/
type=Ind;
run;quit;

/*Weighted Cox PH with Weight 1
(Crossover probabilities)*/
proc genmod data=PFS_Weight
descending;
class PatientID Albumin_cat
MMStage_common creatclear_cat
B2Microglob_cat
Treat_A2(ref=first)/param=ref;
model Event = Treat_A2 Time2 Time3
Time4 Time5 Time6 LOT Albumin_cat
Gender MMStage_common
creatclear_cat B2Microglob_cat/
link=logit dist=bin;
scwgt w1;
repeated subject=PatientID/
type=Ind;
run;quit;

/*Weighted Cox PH with Weight 2
(Crossover probabilities)*/
proc genmod data=PFS_Weight
descending;
class PatientID Albumin_cat
MMStage_common creatclear_cat
B2Microglob_cat
Treat_A2(ref=first)/param=ref;
model Event = Treat_A2 Time2 Time3
Time4 Time5 Time6 LOT Albumin_cat
Gender MMStage_common
creatclear_cat B2Microglob_cat/
link=logit dist=bin;
scwgt w2;
repeated subject=PatientID/
type=Ind;
run;quit;

/*Weighted Cox PH with Weight 3
(Crossover probabilities)*/
proc genmod data=PFS_Weight
descending;
class PatientID Albumin_cat
MMStage_common creatclear_cat
B2Microglob_cat
Treat_A2(ref=first)/param=ref;
model Event = Treat_A2 Time2 Time3
Time4 Time5 Time6 LOT Albumin_cat
Gender MMStage_common
creatclear_cat B2Microglob_cat/
link=logit dist=bin;
scwgt w3;
repeated subject=PatientID/
type=Ind;
run;quit;

/*Weighted Cox PH with Weight 4
(Crossover probabilities)*/
proc genmod data=PFS_Weight
descending;
class PatientID Albumin_cat
MMStage_common creatclear_cat
B2Microglob_cat
Treat_A2(ref=first)/param=ref;
model Event = Treat_A2 Time2 Time3
Time4 Time5 Time6 LOT Albumin_cat
Gender MMStage_common
creatclear_cat B2Microglob_cat/
link=logit dist=bin;
scwgt w4;
repeated subject=PatientID/
type=Ind;
run;quit;
```

R Codes for Plots

Kaplan-Meire Plots

```
par(mfrow=c(2,1))

library(survival)
Surv(survival.data$time_to_death,
survival.data$status_death)
km.estimates<-
survfit(Surv(time_to_death,status_death)~A,
type="kaplan-meier", conf.type="log-log",
data=survival.data)

plot(km.estimates, conf.int=F, ylab="Survival
Probabilities",xlab="Time (Days)",
    main="Survival curves by treatment groups
OS",lwd=2,cex.main=0.9,cex.lab=0.9,cex=0.
5,lty=1:2,col=c("blue","red"))
legend(100, 0.52, lty=1:2,
lwd=2,col=c("blue","red"),
c("MonoTherapy","Combination
Therapy"),bty="n")
abline(h=0.5)

km.estimates1<-
survfit(Surv(time_to_PFS,status_PFS)~A,
type="kaplan-meier", conf.type="log-log",
data=survival.data)

plot(km.estimates1, conf.int=F, ylab="Survival
Probabilities",xlab="Time (Days)",
    main="Survival curves by treatment groups
PFS",lwd=2,cex.main=0.9,cex.lab=0.9,cex=0.
5,lty=1:2,col=c("blue","red"))
```

```
legend(900, 0.988, lty=1:2,
lwd=2,col=c("blue","red"),
c("MonoTherapy","Combination
Therapy"),bty="n")
abline(h=0.5)
```

#Forest plots for each end point
#OS end point

```
method.OS<-c("OS","Ordinary, Propensity
Score and Unweighted
Models","Empty","Ordinary Cox PH(Baseline
Treatment)","Ordinary Cox PH(Baseline
covariates and Treatment)","Cox PH(Stratified
Propensity Score)","Weighted Cox PH(IPTW on
Propensity Score)","UnWeighted
Logistic(Baseline Treatment)","UnWeighted Cox
PH(Treatment History)",
        "Empty","Censoring and Weighted Cox
PH Models for
crossover","Empty","Censoring(Crossover)","Wei
ghted Cox PH with IPTW(Non-Stab
Weight)","Weighted Cox PH with IPCW(Non-
Stab Weight)","Weighted Cox PH
Model1(Weight 1)","Weighted Cox PH
Model2(Weight 2)","Weighted Cox PH
Model3(Weight 3)","Weighted Cox PH
Model4(Weight 4)")

new.OS<-cbind(est.OS,method.OS)

library(rmeta)
forestplot(new.OS,Mean,Lower,Upper,zero=0,
clip=c(-2,2), xlog=TRUE,xlab="Hazard Ratio",

col=meta.colors(box="royalblue",line="darkbl
ue",zero="black"))
```

#PFS end point

```
method.PFS<-c("PFS","Ordinary, Propensity
Score and Unweighted
Models","Empty","Ordinary Cox PH(Baseline
Treatment)","Ordinary Cox PH(Baseline
covariates and Treatment)","Cox PH(Stratified
Propensity Score)","Weighted Cox PH(IPTW on
Propensity Score)","UnWeighted
Logistic(Baseline Treatment)","UnWeighted Cox
PH(Treatment History)",
        "Empty","Censoring and Weighted
Cox PH Models for
Crossover","Empty","Censoring(Crossover)","We
ighted Cox PH with IPTW(Non-Stab
Weight)","Weighted Cox PH with IPCW(Non-
Stab Weight)","Weighted Cox PH
Model1(Weight 1)","Weighted Cox PH
Model2(Weight 2)","Weighted Cox PH
Model3(Weight 3)","Weighted Cox PH
Model4(Weight 4)")

new.PFS<-cbind(est.PFS,method.PFS)

forestplot(new.PFS,Mean,Lower,Upper,zero=0
,clip=c(-2,2), xlog=TRUE,xlab="Hazard
Ratio",

col=meta.colors(box="royalblue",line="darkbl
ue",zero="black"))
```

##Plotting a box plot for estimated propensity scores by treat

```
is.factor(ps_score$Treat)
is.numeric(ps_score$Treat)
treat.f<-factor(ps_score$Treat,
labels=c("Mono","Combination"))
ps.score<-cbind(ps_score,treat.f)

boxplot(P_score ~ treat.f, data = ps.score,
      pars = list(boxwex = 0.3, staplewex =
0.5, outwex = 0.5),
      col = "royalblue",
      main = "Estimated propensity score by
treatment groups",
      xlab = "Treatment group",
      ylab = "Propensity Score",
      ylim = c(0, 1),cex.main=1.0, cex.lab=1.0,
cex.axis=0.9)

boxplot(P_score ~ stratum, data = ps.score,
      pars = list(boxwex = 0.3, staplewex =
0.5, outwex = 0.5),
      col = "royalblue",
      main = "Estimated Propensity score by
strata",
      xlab = "Strata",
      ylab = "Propensity Score",
      ylim = c(0, 1),cex.main=1.2, cex.lab=1.2,
cex.axis=1.2)
```

##IPTW weights for Cox PH Model

```
par(mfrow=c(2,2))
```

##OS end point

```
boxplot(log_weight ~ Time, data = os.IPTW,
      pars = list(boxwex = 0.3, staplewex =
0.5, outwex = 0.5),
      col = "royalblue",
      main = "Estimated IPTW for OS",
      xlab = "Time Interval",
      ylab = "Log Weight",
      ylim = c(-7.5, 3),cex.main=0.8,
cex.lab=0.9, cex.axis=1.0)
```

##PFS end point

```
boxplot(log_weight ~ Time, data = pfs.IPTW,
```

```
      pars = list(boxwex = 0.3, staplewex =
0.5, outwex = 0.5),
      col = "royalblue",
      main = "Estimated IPTW for PFS",
      xlab = "Time Interval",
      ylab = "Log Weight",
      ylim = c(-4, 3),cex.main=0.8,
cex.lab=0.9, cex.axis=1.0)

##IPCW for weighted cox
##OS end point
boxplot(log_weight ~ Time, data = os.IPCW,
      pars = list(boxwex = 0.3, staplewex =
0.5, outwex = 0.5),
      col = "royalblue",
      main = "Estimated IPCW for OS",
      xlab = "Time Interval",
      ylab = "Log Weight",
      ylim = c(0, 8),cex.main=0.8,
cex.lab=0.9, cex.axis=1.0)

##PFS end point
boxplot(log_weight ~ Time, data = pfs.IPCW,
      pars = list(boxwex = 0.3, staplewex =
0.5, outwex = 0.5),
      col = "royalblue",
      main = "Estimated IPCW for PFS",
      xlab = "Time Interval",
      ylab = "Log Weight",
      ylim = c(0, 7),cex.main=0.8,
cex.lab=0.9, cex.axis=1.0)

##Ploting Crossover Probabilities
par(mfrow=c(2,2))

for(i in 1:4){
  interaction.plot(Time[PatientNumber==i],
Prob[PatientNumber==i],
Prob_Cross[PatientNumber==i],
            type="l",col=c(1,2,3,4),legend=F,
ylim=c(0, 1),
            xlab="Time Interval",
ylab="Crossover Probability",lwd=3,
            lty=1, las=1,
main=paste("PatientNumber ", i))


legend("bottomright",c("Prob1","Prob2","Prob3"
,"Prob4"),bty="n",lty=1,lwd=3,

col=c(1,2,3,4),title="Probabilities")
        }
```

Auteursrechtelijke overeenkomst

Ik/wij verlenen het wereldwijde auteursrecht voor de ingediende eindverhandeling:
Adjusting for crossover bias in an observational study for patients with multiple myeloma

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,

Kalinjuma, Aneth Vedastus

Datum: **14/09/2012**