2013•2014
FACULTY OF SCIENCES
Master of Statistics: Biostatistics

Masterproef

Estimate of treatment effect when the treatment changes in time.

Promotor :
Prof. dr. Geert MOLENBERGHS

Promotor :
BART GERRITSE

Gilbert Rukundo

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

Transnational University Limburg is a unique collaboration of two universities in two countries: the University of Hasselt and Maastricht University.

universiteit hasselt
KNOWLEDGE IN ACTION

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

universiteit hasselt | Maastricht University

2013•2014
FACULTY OF SCIENCES
Master of Statistics: Biostatistics

Masterproef

Estimate of treatment effect when the treatment changes in time.

Promotor :
Prof. dr. Geert MOLENBERGHS

Promotor :
BART GERRITSE

Gilbert Rukundo

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

universiteit hasselt | Maastricht University

Estimates of the Treatment Effect when the Treatment Changes in Time

Masters Thesis

October 21, 2013

.

In memory of my Mom & Dad.

Acknowledgment

The success and final outcome of this project required a lot of guidance and assistance from many people and I am extremely fortunate to have got this all along the completion of my project work. Whatever I have done is due to such guidance and assistance and I would not forget to thank them.

I would like to express my deepest appreciation to my internal supervisor, Prof. dr. Geert Molenberghs who was abundantly helpful and offered invaluable assistance, support and guidance. I owe my profound gratitude to my external supervisors Dr. Bart Gerritse and Dr. Teena Bonizzi from Medtronic Bakken Research Center (Maastricht), who guided me all along till the completion of this project work by providing all the necessary information and explanation for developing a good system.

I recognize the financial support from the Flemish Inter-university Council(VLIR) which has enabled me to be among those that benefited from the scholarship grant to pursue this valuable Masters program. I am thankful to constant encouragement, support and guidance from all teaching staffs at Center for Statistics for imparting part of their statistical knowledge on me. I also thank my graduates colleagues (Sept 2013), friends as well as the Uhasselt community for their friendship, love, advice, encouragement, and support during the past two years.

Thanks God, for every blessing you gave me.

Gilbert Rukundo

University of Hasselt, Belgium, October 2013.

Contents

List of Figures

List of Tables

Abbreviations

AAIR: Adaptive Atria Pacing

AF: Atria Fibrillation

AT: Atria Tachycardia

AV: Atria Ventricular

DDDR: Rate Modulated Dual-chamber Pacing

GEE: Generalized Estimating Equation

ICD: Implantable Cardioverter-Defibrillator

IPCW: Inverse Probability of Censoring Weights

ITT: Intention-to-Treat

LVEF: Left Ventricular Ejection Fraction

MVP: Managed Ventricular Pacing

NYHA: New York Heart Association

PP: Per-Protocol

PreFER: Prefer for Elective Replacement

RV: Right Ventricular

Abstract

Estimation of treatment effect in observational studies is not straightforward due to lack of randomization. As a result blind comparison of treatment effects will lead to biased estimates. In this study, treatment switch was a main challenge, though some patients switched treatments. Therefore an analysis that does not account for treatment switch may result into biased estimates. The aim of this study was to estimate the difference in proportion of patients who experienced Atrial Fibrillation (AF) or Atria Tachycardia (AT) with Managed Ventricular Pacing (MVP) programmed ON compared to patients with MVP OFF and common device programming where the endpoint was the time to first event of a persistent AF/AT.

Different methods were proposed for estimating treatment effect in observational studies. For this study, Intention-to-Treat (ITT), Per-Protocol (PP), and Inverse Probability of Censoring weights (IPCW) methods were applied. ITT analysis maintains balance generated from the original random treatment allocation by ignoring anything which happens after randomization. PP restricts analysis to patients who adhered to their assigned treatment in a randomized trial. For these two analyses an ordinary Cox proportional hazard model was fitted in order to estimate the treatment effect. Finally, IPCW was fitted as a most plausible way of estimating treatment effect in presence of treatment switch bias, this method adjusts for bias by creating a pseudo population that would be studied had no selective switch occurred.

The Kaplan-Meier curve shows steep increase in the incidence of persistent AF/AT in the MVP ON arm, when compared to MVP OFF arm. The difference in persistent AF/AT development starts early after enrollment, but remains the same during the follow-up period. The ITT model and PP analysis where 127 patients who switched treatment were excluded in the analysis lead to the same conclusion of no treatment effect ($p = 0.07$ and $p = 0.14$), adjusting for age and gender. Although method like IPCW method correct for confounding induced by the treatment switch, this model also fail to detect a significant effect of the treatment over time ($p = 0.22$).

For all the considered methods, the difference between the two randomization arms during the 2 years follow-up period is not statistically significant. Therefore, based upon these data, MVP ON is not shown to be superior to MVP OFF and common device programing in terms of freedom from persistent AF/AT.

CHAPTER 1

INTRODUCTION

1.1 Problem Description

The Medtronic Bakken Research Center in Maastricht, the Netherlands, supplied us with the data of a longitudinal observational study (PreFER MVP), in which patients were assigned randomly to one of the two programming modes of an implantable device. However, patients were allowed to switch to a different treatment (non-randomized) for clinical and technical reasons. This presents a challenge since programming modes comparison cannot be accounted by simple statistical methods. As a result, comparing treatments in this nature of data may simply reflect the underlying difference between treatment groups and not treatment effect (Curtis et al., 2007). Estimating the effect of a programming mode presents challenges; including the need to address switching program and time-dependent confounders (Faries et al., 2010; Hernan et al., 2000, 2004). Therefore, it is difficult to know the actual effect of a certain programming mode if these challenges are not controlled for during estimation of programming mode effect. The aim of this thesis is to:

- Estimate the difference in proportion of patients experienced Atrial Fibrillation (AF) or Atria Tachycardia (AT) with Managed Ventricular Pacing (MVP) programmed ON compared to patients with MVP OFF and common device programming where the endpoint was the time to first event of a persistent AF/AT.

1.2 Medical Background and MVP Feature

A number of clinical studies (Andersen et al., 1997; Nielsen et al., 1998) over the past few years have shown that, in patients with intact Atria Ventricular (AV) conduction, unnecessary chronic right ventricular (RV) pacing can cause a variety of detrimental effects, including Atrial Fibrillation (AF) or Atria Tachycardia (AT). These effects are believed to result from the mechanical dysynchrony and ventricular chamber dysfunction that occurs with chronic, single-site, apical ventricular stimulation.

Therefore a new pacing modality, Managed Ventricular Pacing (MVP), was designed to give preference to natural heart rythm by minimizing right ventricular pacing. This is accomplished by automatically switching between single chamber atria and dual-chamber pacing based on specific patient needs. MVP is an atrial-based dual chamber pacing mode that provides functional AAI/R pacing with ventricular monitoring and back-up DDD/R pacing only as needed during episodes of AV block.

The reversibility of the detrimental effect caused by ventricular pacing has been initially investigated in small patient populations with short pacing duration in AAI and needs further investigation. MVP was developed to address the inherent limitations of AAI/R pacing and dual chamber modes that incorporate a fixed or unnecessary right ventricular pacing. MVP provides atrial based pacing with ventricular backup. In case of loss of AV conduction the device switches to DDDR or DDD mode. Periodic checks are performed, and if AV conduction resumes, the device switches back to AAIR or AAI mode. The Medtronic ICD and pacemaker models used in this study are all equipped with the MVP feature (Garutti & Vainer, 2012).

1.3 Challenges in the Analysis

One of the main difficulties in estimating the effect of the treatment in randomized observational studies is the confounding caused by treatment switch. It is common to allow patients from one programming mode to switch to the other programming mode if necessary and for clinical or technical reasons. Thus, the overall advantage associated with the experimental programming mode cannot be estimated with confidence based on the Intention-to-Treat (ITT) data, because a proportion of the patients randomized to the control programming mode will have shifted to the experimental programming mode and vice-versa. Equally a subset analysis approach where

patients who switched program are excluded from the analysis at the time of switching is also likely to be confounded because program switching is unlikely to have occurred at random. In such circumstances censoring is informative and the randomization of the trial is compromised. Treatment switch has been an important issue in the analysis of clinical trials but the methods used to account for the impact switching on the treatment effect have generally been simplistic. Most regularly censoring approaches have been used but often the switch has been ignored and standard ITT analysis is conducted which produce heavily biased results. More recently, statistical techniques to address the switching treatment problem have been developed.

1.4 Outline

In the second chapter the data are introduced. This includes: the population and variable description followed by a summary of the patients' baseline characteristics. In the third chapter, we address the methods which will be used for analysis. These methods include: Intention-to-treat (ITT) which analyses the data based only on their initial assignment. Subsets analysis are done where information from patients who switched program mode are excluded from the analysis, as well as unweighted Generalized Estimating Equation (GEE) and weighted Cox Proportional Hazard Models where the basic idea is to weight the data by the Inverse Probability of Censoring (=switching) Weights (IPCW). In Chapter 4, we present the results obtained by using the above mentioned methods followed by a comparison between them. Finally, Chapter 5 provides a summary discussion of the results of this study and ends with limitations and recommendations for further research.

CHAPTER 2

DATA DESCRIPTION

2.1 Population and Variable Description

The data set used in this study comes from an observational longitudinal study (PreFER MVP) with a survival endpoint, namely time to first event of persistent AF/AT. It consists of 605 patients who were enrolled from 75 different medical sites around the globe, including: Europe, Canada, the Middle East and Australia. All the considered patients had an implantable device (Pacemaker or ICD), equipped with MVP feature. Patients were randomized into two different groups; experimental group and control group with MVP programmed ON and MVP programmed OFF respectively.

The follow-up time was a duration of two years, starting from 2006 where measurements were recorded by the device on a daily basis. These patients could continue in the study until they experienced persistent AF/AT or if they met any of the exclusion criteria; see Garutti and Vainer (2012) for details. The time to persistent AF/AT will be used as response variable for estimating the effect of MVP program modes.

Several variables were considered in this study. These variables include: the response variable: time to first event of persistent AF/AT and baseline covariates. Table 2.1 presents a full list of variables considered in the analysis.

Table 2.1: *Variable description*

Variable	**Description**	**Code/Value**
PatID	Patient Identification number	
MVP^*	Actual MVP programming	0=MVPoff, 1=MVPon
Randomization	Randomization	0=MVPoff, 1=MVPon
Day	Day from Implant	days
DurSE2	Time to persistent AF/AT or last visit days	days
SE2	Had persistent AF/AT	0=no, 1=yes
GEN	Gender	0=Male, 1=Female
Age	Age of the patient years	years
Country	Country(ies) of origin	Name(character)

*=*Time-varying*

2.2 Patients Characteristics and Risk Factors by Outcome

Descriptive statistics were used to give a first insight into the distribution of the variables. Summaries of patients' characteristics in the treatment groups are shown in Table 2.2. There were no missing data for the covariates age and gender, two patients had only one observation each and another 14 had missing observations on the response variable. A total of 605 patients were enrolled in the trial with almost equal number of patients over the treatment groups and almost equal number of males and females by treatment program. The number of events for the response of interest was a bit higher in patients who were in the MVP programmed ON arm (14%) as compared to MVP programmed OFF arm (10%). There was a remarkable difference in median follow-up time to persistent AF/AT among the two treatments program modes, i.e., MVP ON (397 days) and MVP OFF (487 days), with the over all median follow-up period of 441 days for total cohort. More than 75% of the patients considered in this study had more than 75 years of age at the start of the study. There was no big difference in the number of patients who switched treatments and no difference in the number of patients who had persistent AF/AT in the two treatment arms respectively.

Table 2.2: *Patients characteristics (N=605)*

Baseline			
	MVP ON	**MVP OFF**	
Variable	number	number	Total
Number of Patients	299	306	605
Gender			
Males (%)	174 (58.20)	191 (62.42)	365 (60.33)
Female (%)	125 (41.80)	115 (37.58)	240 (39.67)
Median Age	76.92	75.49	76.74
Median DurSE2	397	487	441
After 2 years follow-up period			
Switched Program (%)	71 (23.75)	56 (18.30)	127 (20.99)
Persistent AF/AT (%)	41 (13.71)	30 (9.80)	71 (11.74)

2.3 Software Used

The statistical analysis was performed using SAS version 9.3. All statistical tests were performed at 5% level of significance unless stated otherwise.

CHAPTER 3

METHODOLOGY

3.1 Statistical Analysis

Various methods have been proposed for analysis in the literature and situations where patients depart from their randomized treatment (Morden et al., 2011; Rimawi & Hilsenbeck, 2012). We refer to them here as simple methods, which tend to involve only small adjustments to standard survival techniques. This section will focus on three of these: intention-to-treat, per-protocol analysis, which excludes patients if they switched treatment, and Inverse Probability of Censoring Weights (IPCW). Complex observation-based methods such as Marginal Structural Models (MSM) also exist (Faries et al., 2010; Westreich et al., 2010). For this analysis, the later method (MSM) will not be considered here due to lack of time dependent covariates with sufficient observations. Methods not accounting for treatment switch are discussed before proceeding to IPCW method. They represent simple techniques that have often been used to analyze data in which a treatment switch has occurred. These methods were included in the analysis to allow comparisons with IPCW method. The original data set had a survival data structure with one observation per subject. Due to the nature of the weights to be estimated (time specific) standard software for fitting Cox regression model cannot be used for this type of analysis. In order to fit the weighted Cox proportion hazard model with this type of weights, survival data structure needs to be transformed into panel (longitudinal) data (Hernan et al., 2000). In this study, data transformation was done based on the observed time to event.

3.2 Intention-to-Treat

Randomized trials often suffer from different kinds of complication. These include: noncompliance and missing outcomes. One potential solution to this problem is a statistical concept called Intention-to-Treat (ITT) analysis (Fergusson et al., 2002). ITT analysis includes every subject who is randomized according to randomized treatment assignment, it ignores noncompliance, protocol deviations, withdrawal, and anything that happens after randomization. ITT analysis maintains balance generated from the original random treatment allocation. In this analysis method, the estimate of treatment effect is generally conservative and the obtained results from an ITT population in a superiority trial tend to shift toward the null hypothesis and reduce chances of false-positive conclusions (Type I error) (White, 2012).
In this study, an intention-to-treat analysis was done by fitting a multivariate Cox proportional hazards regression model of the form:

$$h_i(t|X_i, Trt_i) = h_o(t) \exp(\beta X_i + \gamma Trt_i), \tag{3.1}$$

where:

- $h_0(t)$ is the baseline hazard function of a patient in the control group when all the covariates (X_i) are zero;
- Trt_i is the treatment (0= MVPOFF , 1= MVPON);
- X is the vector of regression parameters and $exp(\beta)$ is the hazard ratio (HR);
- γ is the log hazard for the treatment.

The Cox proportional hazard model, which is a regression method for survival data, provides an estimate of the hazard ratio and its confidence interval. The hazard ratio is an estimate of the ratio of the hazard rate in the treated versus the control arm. The hazard rate is the probability that if the event in question has not already occurred, it will occur in the next time interval, divided by the length of that interval. The time interval is made very short, so that in effect; the hazard rate represents an instantaneous rate. An assumption of proportional hazards regression is that the hazard ratio is constant over time (Spruance et al., 2004).

3.3 Per-protocol Analysis

When patients deviate from their randomized treatment, an analysis restricted to patients who adhered to their assigned treatment in a randomized trial (omitting patients who dropped out of the study or switched treatment for any reason) might be performed. Though a per-protocol analysis may be appropriate in some settings, it should be properly labeled as a non-randomized observational comparison and any exclusion of patients from the analysis compromises the randomization may lead to bias in the results since all patients randomized are no longer included. This raises concerns about whether important unknown factors that influence outcome are equally distributed across comparison arms. In theory, random switch of treatment would tend to have no effect, preserving the true effects of the treatment but problems arise when the switching is not random. The per-protocol analysis was done by fitting a multivariate Cox proportional hazard regression model (3.1). However, as Morden et al. (2011) noted, this analysis should not be done alone, it has to be compared to other methods.

3.4 Unweighted Marginal Model (GEE)

Two Generalized Estimating Equations (GEE) (Zeger & Liang, 1986; Liang & Zeger, 1986) models which use treatment as a time varying covariate were considered, these include: a model with outcome as had an event (AT/AF) or not given baseline covariates and randomized treatment and a model with outcome as had an event (AT/AF) or no event given baseline covariates and treatment history. The term marginal in this context indicates that the model for the mean response depends only on the covariates of interest, and not on any random effects or other responses. The marginal model is used when the researcher investigates the overall population average trend as a function of the covariates while accounting for the correlations in the data. The association structure is then typically captured using a set of association parameters, such as correlations, odds ratios, etc (Molenberghs & Verbeke, 2005).

Suppose that Y_{ij} is a binary response, taking the value of 0 denoting failure (no AF/AT) or 1 denoting success (had AF/AT), and it is of interest to relate change in $E(Y_{ij}) = Pr(Y_{ij} = 1)$ to the covariate. With binary response, the distribution of each Y_{ij} is Bernoulli and the probability of success is often modeled using a logit link function. The marginal expectation of the response $E(Y_{ij}) = \mu_{ij}$ depends on the covariate X_{ij} through a know link function $g(\mu_{ij})=X'_{ij}\beta$. The GEE

estimator of β for marginal models can be thought of as arising from minimizing the following objective function (Liang & Zeger, 1986).

$$\sum_{i=1}^{N}\{Y_i - \mu_i(\beta)\}' V_i^{-1}\{Y_i - \mu_i(\beta)\}, \tag{3.2}$$

with respect to β where $V_i = A_i^{\frac{1}{2}}\text{Corr}(Y_i)A^{\frac{1}{2}}$ is the marginal covariance matrix of Y_i which is treated as known (by ignoring its dependence on β through μ_i), $\mu_i = \mu_i(\beta) = X_i\beta$ is the vector of mean response and $\text{Corr}(Y_i)$ is the marginal correlation matrix. Using calculus, it can be shown that if the minimum of the function given by (3.2) exists, then, the regression parameters β are estimated by solving the estimating equation:

$$\sum_{i=1}^{N}\frac{\partial \mu_{ij}}{\partial \beta_k} V_i^{-1}\{Y_i - \mu_i(\beta)\} = 0. \tag{3.3}$$

In this study, these two models will be used to allow a comparison with a weighted model.

3.5 Weighted Cox Proportional Hazard Model with Inverse Probability of Censoring Weight(IPCW)

Some arguments against ITT analysis appear valid. To begin with, if a subject who actually did not receive any treatment is included as a subject who received treatment, then it indicates very little about the efficacy of the treatment. Also, heterogeneity might be introduced if noncompliants, dropouts, and compliant subjects are mixed together in the analysis and for this reason, interpretation might become difficult if a large proportion of participants switched to the opposite treatment arm. A more used alternative method is the Inverse Probability of Censoring (=Switching) Weights (IPCW). This method adjusts for bias associated with time-dependent confounders that are affected by prior treatment or exposure (e.g., dropout due to adverse effects). However, this method has also been applied to control for selective switch which is more likely among high risk subjects (Cain & Cole, 2009). IPCW analysis attempts to create a pseudo population that would be studied had no selective switch occurred (Rimawi & Hilsenbeck, 2012). This pseudo population is created by weighting each not artificially censored patient's contribution to a given risk set. Specifically at time $t_{(i)}$, each patient is assigned weight of $W_{t_{(i)}}$ that is inversely proportional to the estimated conditional probability that the patient remained not artificially censored through time $t_{(i)}$. The conditional probability and weight

$W_{(t_{(i)})}$ are estimated by fitting a discrete-time pooled logistic regression model for artificial censoring, in which the common predictors of the endpoint of interest and the artificial censoring mechanism are included as covariates in the model.

$$\text{logit}P(C_{ij} = 0|X_i) = \beta_0 + \beta X_i \tag{3.4}$$

where: C_{ij} represents status of switching for i^{th} patient at j^{th} time point with $C_{ij} = 0$ for patients who switched treatment, whereas X_i represents a vector of covariates for patient i. The ability of the IPCW method to create the pseudo population that would exist in the absence of artificial censoring depends on whether the assumptions of exchangeability and correct model specification are met. Estimated weights that are extreme in value indicate model misspecification or nonpositivity. In turn, an estimate of the survival function based on such weights may fail to correct for the bias introduced by the treatment switch. Whether the weights are extreme because of model misspecification or non-positivity cannot be known with certainty (Howe et al., 2011).

CHAPTER 4

RESULTS

4.1 Exploratory Data Analysis

4.1.1 Treatment Switch Statistics

The dataset used in thi study, combined the information obtained from 605 patients, these patients were randomized into two treatment arms: MVP OFF ($n = 305$) and MVP ON ($n = 299$). After a follow-up period of two years, only 127 (21.49%) patients had deviated from their initial treatment with 71 (12.01%) patients switched treatment to MVP ON from MVP OFF and 56 (9.48%) patients switched to MVP OFF from MVP ON. Switching treatment had been considered to confound the interpretation of long-term follow-up data and raise the issue of how to deal with the deviation from randomized treatment in general. Table 4.1 shows the proportion of patients who switched from one treatment to another. This table exhibits that there is a considerable difference in the number of patients who experienced persistent AF/AT between the two switching modes, i.e., 13 (18.30%) for patients who switched and 58 (81.70%) for patients who did not switch. Almost a quarter of all patients switched treatment and one would wonder if this number is enough to alternate the results which would be obtained if this treatment switch is ignored (ITT).

Table 4.1: *Treatment switch descriptive statistics*

	Randomized to MVP OFF		Randomized to MVP ON		Total	
	Switched	Stayed	Switched	Stayed	Switched	Stayed
Number	$n = 71$	$n = 230$	$n = 56$	$n = 234$	$n = 127$	$n = 464$
Gender						
Male	43	146	32	135	75	281
Female	28	84	24	99	52	183
Median age	75.39	77.48	81.33	77.38	77.24	77.46
Median DurSE2	682	412	354	440	632	430
Persistent AF/AT	7	23	6	35	13	58

Missings = 14

4.1.2 Kaplan-Meier Curve

Kaplan-Meier estimate is one of the best options to be used to measure the fraction of subjects living for a certain amount of time after treatment (Goel et al., 2010). In clinical trials, the effect of an intervention is assessed by measuring the number of patients who have not experienced the event of interest after that intervention over a period of time. Despite difficulties associated with patients or situation, the Kaplan-Meier estimate is the simplest way of computing the survival/incidence over time. One of its advantages is that it requires only very weak assumptions and utilizes the information content of both: full observed and right censored data. In this study, the Kaplan-Meier curve was used in order to get insight into the shape of the incidence function for each treatment arm. Figure 4.2 shows that there seems to be an increase in the incidence of persistent AT/AF in the MVP ON arm, when compared to the MVP OFF arm. The difference in persistent AF/AT development starts early after enrollment, but remains the same during the follow-up period.

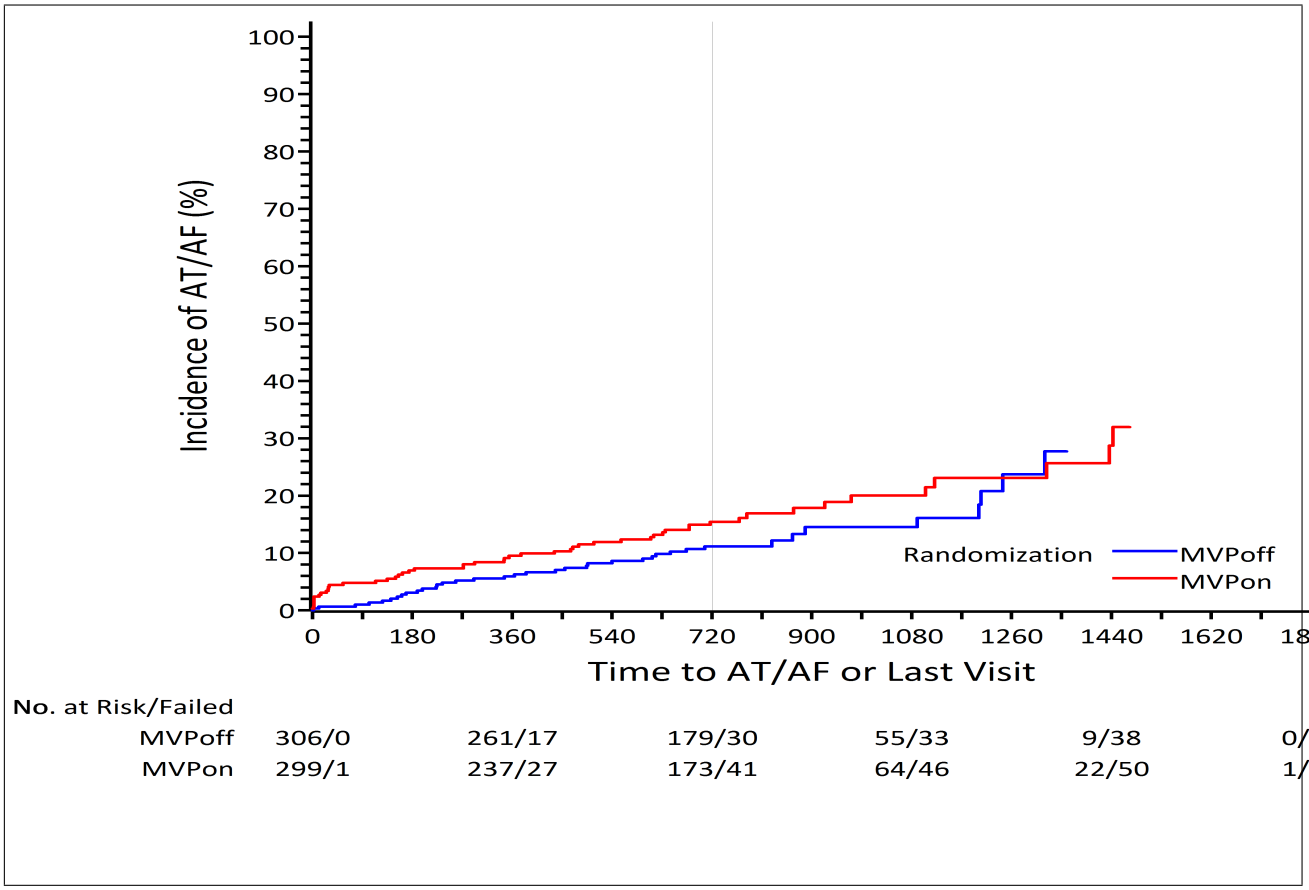

Figure 4.1: *Incidence of persistent AF/AT*

MVP ON arm showed a steep increase in persistent AT/AF development directly after enrollment, which may be explained by the fact that AT/AF was not documented/known at baseline, but immediately as such documented in the new device and patients were having a device replaced. This means that, patients may be more predisposed to tolerating ventricular pacing, so when we implant device which minimize ventricular pacing; maybe this causes some remodeling or change in the heart. Note that, in both endpoints (2 years follow up period) incidence curves were not touching each other. Moreover these plots also showed no lack of proportionality in estimated curves, which suggests a validation of proportional hazards assumption.

4.2 Statistical Analysis

4.2.1 Cox Proportional Hazard Model with Treatment at Baseline as the Only Covariate

Whereas the Kaplan-Meier method with log-rank test is useful for comparing survival curves in two or more groups, Cox proportional hazards regression model allows analyzing the effect of risk factors on survival. The most naive model that only includes treatment at randomization shows that the difference between the two randomization arms is not statistically significant for neither considered follow-up period (one year and two years; $p = 0.133$ and $p = 0.078$). Therefore, based on the given data and the method used so far, the MVP ON does not appear to be superior to MVP OFF and common clinical device programing in terms of freedom from AF/AT. The Table 4.2 below presents the hazard ratio together with the confidence interval associated to the model with randomization as the only covariate.

Table 4.2: *Cox PH model estimate for time to AF/AT by randomized arm. In the hazard ratio the numerator is the hazard rate of MVP ON and the denominator is the hazard ratio of MVPOFF*

Parameter	**Pr> ChiSq**	**Hazard Ratio**	**95% Hazard Radio C.I**
	After one year follow up period		
Randomization	0.133	1.564	[0.873 ; 2.800]
	After two years follow-up period		
Randomization	0.078	1.521	[0.954 ; 2.424]

4.2.2 Cox Proportional Hazard Model with Age, Gender, and Randomization as Covariates

Intention-to-treat (ITT) Analysis

In the ITT analysis, using a standard Cox proportional hazards model with age, gender, and treatment at randomization as covariates, the hazard ratio shows that female patients are less likely to experience persistent AF/AT, adjusting for treatment and age for one year follow up time 0.521 (95% C.I: 0.274;0.989). Results in Table 4.3 indicate that there was no signif-

icant effect of the treatment on neither follow-up time (one year and two years; $p = 0.116$ and $p = 0.058$). While analysis of this type is valid, it may underestimate the appropriate effectiveness of a treatment. For example, if the experimental treatment (MVP ON) truly is superior to the control treatment (MVP OFF), and some patients have switched from control to experiment, and are therefore receiving the benefits of this, using an ITT analysis will make the treatments appear more similar than they really are. The benefit of this type of analysis is that randomization balance between treatment arms is maintained.

Table 4.3: *ITT: Cox PH model estimates for time to AF/AT with age and gender variables and randomization as covariates. In the hazard ratio the numerator is the hazard rate of MVP ON and the denominator is the hazard ratio of MVP OFF*

Parameter	**Pr> ChiSq**	**Hazard Ratio**	**95% Hazard Radio C.I**
	After one year follow up period		
Age	0.230	1.018	[0.989 ; 1.049]
GEN(Female)	0.046	0.521	[0.274 ; 0.989]
Randomization	0.116	1.598	[0.891 ; 2.867]
	After two years follow-up period		
Age	0.285	1.012	[0.990 ; 1.035]
GEN (female)	0.149	0.699	[0.430 ; 1.136]
Randomization	0.070	1.542	[0.966 ; 2.463]

Subset (Per-protocol) Analysis Model

Per-protocol analysis compares treatment arms and includes only those participants who completed the treatment protocol originally allocated. In this paragraph, patients who switched treatment, were excluded from the analysis and results showed a non-significant treatment effect ($p = 0.111$; Table 4.4). However, if done alone, results obtained from this method may lead to bias. By comparing the results from the intention-to-treat and per-protocol analyses, the potential bias of confounding in the estimate of treatment effectiveness indicated by per-protocol analysis can be estimated. The results of the two analyses were similar, which suggests that non-adherence to randomized treatment was limited. Moreover, both models show no lack of proportionality.

Table 4.4: *Subset Analysis: Cox PH model estimates for time to AF/AT with age and gender variable and randomization as covariates. In the hazard ratio the numerator is the hazard rate of MVPON and the denominator is the hazard ratio of MVPOFF*

Parameter	**Pr> ChiSq**	**Hazard Ratio**	**95% Hazard Radio C.I**
	After one year follow up period		
Age	0.310	1.018	[0.983 ; 1.055]
GEN(Female)	0.059	0.481	[0.225 ; 1.028]
Randomization	0.121	1.722	[0..866 ; 3.424]
	After two years follow-up period		
Age	0.216	1.017	[0.990 ; 1.044]
GEN(Female)	0.364	0.780	[0.456 ; 1.334]
Randomization	0.138	1.491	[0.879 ; 2.527]

4.2.3 Unweighted Marginal Model (GEE)

To perform this analysis, given data were transformed into panel data so that the treatment variable could be used as a time varying covariate, this was done based on the variable indicating the time to event (DurSE2). GEE was chosen as an alternative to the above methods (see sections 4.2.1 and 4.2.2) that account for the correlation in the data. For simplicity, we assumed an independence working correlation matrix. This choice is justified since the GEE method is robust against misspecification of the working correlation structure and it provides efficient parameter estimates as longer as the mean model is well specified. The first model (Model1) considered had binary response variable 0 as 'fail' (No AF/AT) or 1 as 'success' (had AF/AT) together with visit, treatment at randomization, age and gender variables as covariates (Table 6.1) whereas the second model (Model2) had treatment history, visit, age and gender variables as covariates. Model1 showed no significant effect associated to the treatment over time ($p = 0.564$). An improvement was made to Model1 where the treatment at randomization was replaced by the received treatment in Model2. However, no treatment effect was observed in this case neither ($p = 0.280$).

4.2.4 Inverse Probability of Censoring (=Switching) Weights (IPCW)

Pooled logistic regression model for switching indicator

Results from the pooled logistic regression model for the probability of switching given past treatment history, age and gender covariates showed that gender and previous treatment (given the day of visit) were higher predictors of the treatment switch ($p < 0.001$ for both; Table6.3). The boxplot of the weights (Figure:4.3) shows a decrease in the mean and the variance of weights over time. Note that patients who did not switch treatment and had similar history as those that switched treatment got bigger weights, i.e., these patients were 'counted' more than once to make up for people like them that were censored due to switch. Thus the reweighted population is no longer a biased sample.

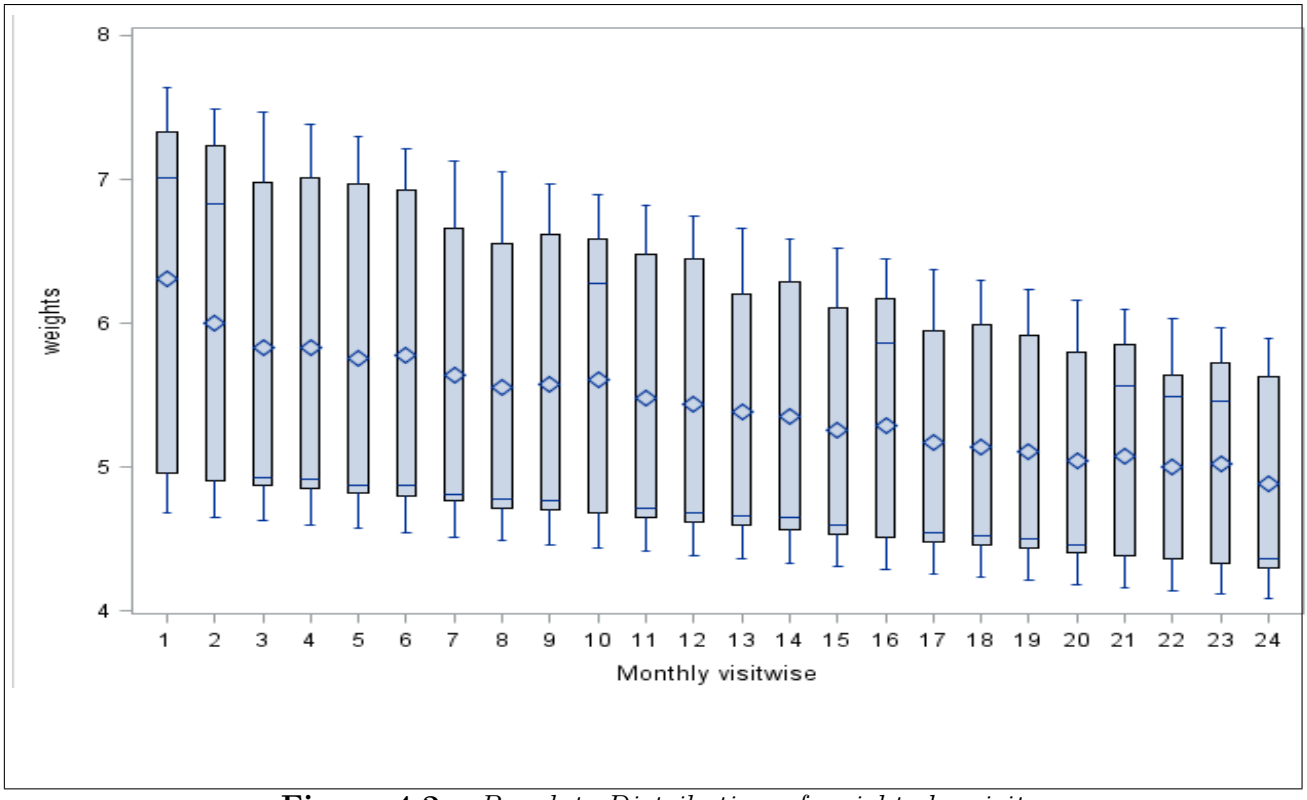

Figure 4.2: *Boxplot: Distribution of weights by visit*

Inverse Probability of Censoring (=Switching) Weights (IPCW)

To adjust for confounding due to treatment switch, the parameter of the treatment effect was estimated by fitting a weighted Cox proportional hazards model where weights were calculated based on the individual probability of switching treatment given age and gender covariates and treatment taken up to time j. The results obtained from this model show no significant effect of the treatment over time. Thus, on average, over the considered time period, the

treatment administrated to a patient does not influence the occurrence of the persistent AF/AT ($p = 0.224$; Table 6.3). On the other side, no significant effect of gender or age observed after a follow-up period of two years. Artificial censoring correction using IPCW might fail when data that violate the exchangeability assumption are used. With the current data, it was not easy to verify this assumption since this would require additional data. Finally, for GEE and weighted Cox proportional hazard fitted models, results from the parameter estimates of interest (VISIT*TRT) were combined in one plot (Figure 4.3) to help for an easy comparison of these models.

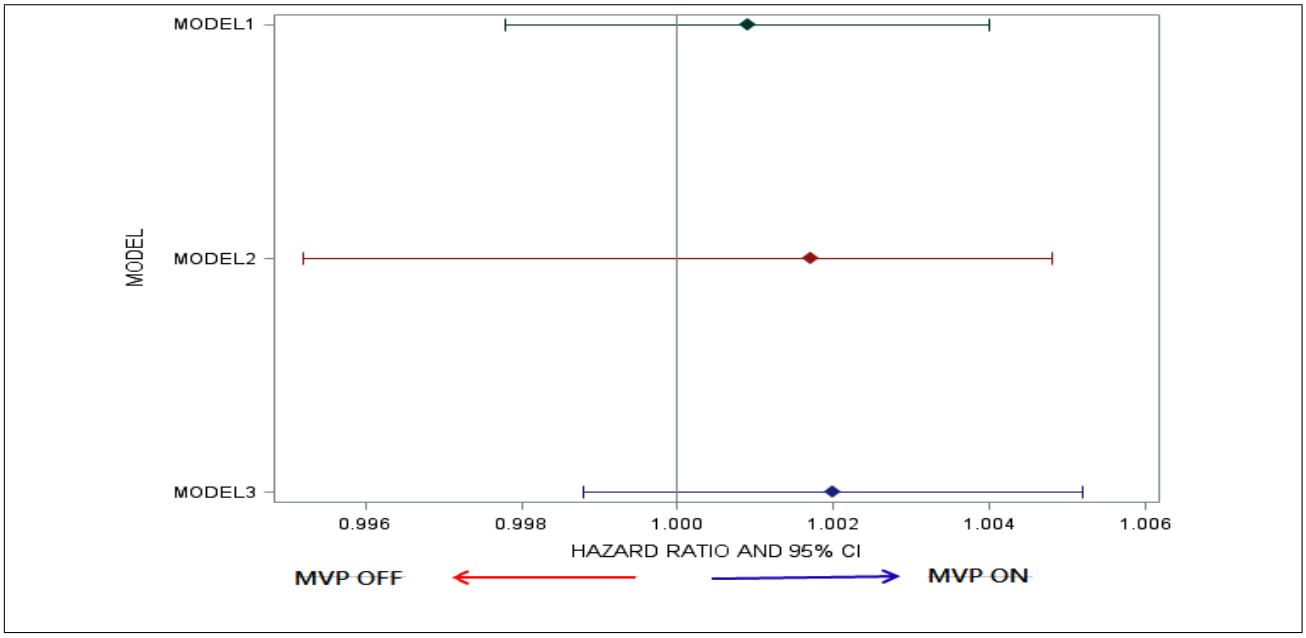

Figure 4.3: *Forest plot, hazard ratios and their 95% CI for VISIT*TRT estimates; Model1: Unweighted GEE Model with outcome as had an event (AT/AF) or no event given age and gender covariates and randomized treatment; Model2: Unweighted GEE model with outcome as had an event(AT/AF) or no event given age and gender covariates and treatment history; Model3: Robust estimates from a weighted Cox PH model with weights based on age and gender covariates and treatment history.*

4.2.5 Comparison of the considered methods

Figure 4.4 combines the hazard ratio and 95% CI for different models. The ITT model which is a Cox proportional hazard model and the subset analysis where 127 patients who switched treatment were excluded in the analysis lead to the same conclusion in both follow-up duration. That is, there is not enough evidence to conclude that MVP ON is superior to MVP OFF in

terms of freedom from persistent AF/AT. This implies that maybe the treatment switch has no big impact on the true treatment effect in this study. Although methods like IPCW correct for confounding induced by the treatment switch, this model failed to detect a significant effect of the treatment over time. It can be noted that all the methods used present a consistent message, i.e., based on the data considered in this study, MVP ON is not shown to be superior to MVP OFF and common clinical device programming in terms of freedom from persistent AF/AT.

CHAPTER 5

DISCUSSION AND CONCLUSION

The objective of this study was to estimate the difference in proportion of patients experienced Atria Fibrillation (AT) or Atria Tachycardia (AT) with managed ventricular pacing (MVP) programmed ON as compared to patients with MVP OFF and common device programming based on a large dataset from a prospective observational study. The main challenge in this study was to compare different kind of the methods, these included, ordinary Cox proportional hazard model, unweighted GEE models and weighted Cox proportional hazard model. To fit the latter two models data were transformed into a longitudinal sequence in order to estimate the treatment effect. Time to event outcome measures are not generally affected by treatment switch in cases where switch is only done after an event had occurred. This was not the case in this study as some of the patients switched treatment before experiencing a persistent AF/AT ($n = 127$). Thus, the survival estimates were adjusted for treatment switch.

From the Kaplan-Meier plot in exploratory data analysis, it is observed that MVP ON arm showed a steep increase in persistent AT/AF development directly after enrollment, which may be explained by the fact that AF/AT was not documented known at baseline, but immediately as such documented in the new device. Whereas the Kaplan-Meier method with log-rank test is useful for comparing survival curves, Cox proportional hazards regression allows to analyze the effect of several risk factors on survival. With an ITT method, the Cox proportional hazard model compares the two treatment arms based on the treatment originally allocated. This is regardless of whether the patient began the treatment allocated, subsequently withdrew from

the trial, did not adhere to the protocol of the allocated treatment, or received a different treatment from that originally allocated. In this trial, an ITT analysis indicated that 30 of 306 (9.8%) patients treated with MVP OFF and common clinical device programming had experienced persistent AF/AT after a follow up duration of two years as compared with 41 of 299 (13.7%) treated with MVP ON. The estimated hazard ratio from this method showed a non significant difference between the two randomization arms. Although ITT analysis preserves the sample size, i.e keep the statistical power, it has been criticized for being too cautious and thus being more susceptible to type II error (Hollis & Campbell, 1999). One most common alternative to ITT analysis is the per-protocol analysis also know as modified ITT analysis which restricts the comparison of the treatments arms to the ideal patients, results from this model agree with those obtained from the ITT analysis, that there seems to be no statistically significant difference between the two randomized arms. Though a per-protocol analysis may be appropriate in some settings, it should be properly labeled as a non-randomized, observational comparison. Any exclusion of patients from the analysis compromises the randomization and may lead to bias in the results.

To adjust for confounding due to treatment switch, a weighted Cox proportional hazard model was applied. The weights were calculated to represent the inverse probability of censoring given factors affecting the treatment switch. Due to treatment switch, the hazard ratios calculated from the models which do not take into account switching are biased. Although not significant, the results obtained from the IPCW model showed a small improvement as compared the preceding models, this is shown by a significant treatment effect and a significant visit effect individually. However, due to the presence of interaction term between visit and treatment (VISIT*TRT) which is not significant, no conclusion can be made from the main effect of these covariates.

As a summary of this discussion, all the considered methods lead to the same conclusion that: based upon these data, MVP ON is not shown to be superior to MVP OFF and common clinical device programming in terms of freedom from persistent AF/AT.

Comparing results of this study to other studies that have used statistical techniques to account for treatment switch; a number of randomized trials have compared common clinical pacemakers and MVP mode. Although some trials have demonstrated that MVP ON mode significantly

reduced the frequency of ventricular pacing in patients with sinus node disease and AV block (Gillis et al., 2006), others have demonstrated only modest (Brignole et al., 2005) or no benefit at all (Sweeney et al., 2003; Albertsen et al., 2008).

Study Limitations

The following limitations can be noted:

- Atria passing percentage was not collected at baseline.
- Missing device data: 25% of the visit Case Report Forms were sent without a save to disk. This affected the analysis of our data since methods which correct for missingness may not work properly in presence of such amount of missing data. Therefore some covariates were excluded from the analysis due to a huge number of missing measurements, this is the case of two time varying covariates: LVEF and NYHA.
- The longer follow up period for the MVP ON randomized patients, as a result of the 2:1 randomization in the first 1.5 year. the mean extra follow up time for MVP ON patients, compared to the MVP OFF randomized patients was 2 months.

References

Albertsen, A. E., Nielsen, J. C., & Poulsen, S. H. (2008). Biventricular pacing preserves left ventricular performance in patients with high-grade atrio-ventricular block: a randomized comparison with ddd(r) pacing in 50 consecutive patients. *Europace*, ***10***(3), 314–320.

Andersen, H., Nielsen, J., & Thomsen, P. (1997). Long-term follow-up of patients from a randomized trial of atrial versus ventricular pacing for sick-sinus syndrome. *Lancet*, ***350***, 1210–1216.

Brignole, M., Gammage, M., & Puggioni, E. (2005). Comparative assessment of right, left, and biventricular pacing in patients with permanent atrial fibrillation. *Eur Heart J.*, ***26***(11), 712–722.

Cain, L. E., & Cole, S. R. (2009). Inverse probability of censoring weights for the correction of time-varying noncompliance in the effect of randomized highly active antiretroviral therapy on incident aids or death. *Statistics in Medicine*, ***28***(12), 1725-1738.

Curtis, L., Hammill, B., Eisenstein, E., Kramer, J., & Anstrom, K. (2007). Using inverse probability-weighted estimators in comparative effectiveness analyses with observational databases. *Medical Care*, ***45***(10). Retrieved from `http://www.hsph.harvard.edu/merckworkshop/files/2013/01/inverse_probability_weighting.pdf`

Faries, D., Leon, A., Haro, J., & Obenchain, R. (2010). *Analysis of longitudinal observational data using marginal structural models* (1st ed.). SAS Institute Inc., Cary, NC, USA: Medical Care.

Fergusson, D., Aaron, S. D., Guyatt, G., & Herbert, P. (2002). Post-randomisation exclusions: the intention to treat principle and excluding patients from analysis. *BMJ*, ***325***, 652–654.

Garutti, C., & Vainer, L. (2012). *Time scale and adjusted survival curves for marginal structural cox models* [Report]. Medtronic Bakken research center, Maastrich.

Gillis, A. M., Purerfellner, H., Carsten, W. I., Stunthorn, H., Kacet, S., M., M. A., ... Boriani, G. (2006). Reducing unnecessary right ventricular pacing with the managed ventricular pacing mode in patients with sinus node disease and av block. *PACE*, ***29***(23), 697–705.

Goel, M. K., Khanna, P., & Kishore, J. (2010). Understanding survival analysis: Kaplan-meier estimate. *Int J Ayurveda Res*, ***1***(4), 274–278.

Hernan, M., Brumback, B., & Robin, J. (2000). Marginal structural models to estimate the

causal effect of zidovudine on the survival of hiv-positive men. *Epidemiology*, *11*(**5**), 561–570. Retrieved from http://www.epiresearch.org/playlists/files/Hernan_EPI_2000_11_561.pdf

Hernan, M., Diaz, S., & Robins, J. (2004). A structural approach to selection bias. *Epidemiology*, ***15***(5), 615–625. Retrieved from http://www.idepi.hsph.harvard.edu/faculty/miguel_hernan/files/hernan_epidemiology04.pdf

Hollis, S., & Campbell, F. (1999). Review what is meant by intention to treat analysis? survey of published randomised controlled trials. *BMJ*, ***319***(7211), 670–674.

Howe, C., Cole, S., Chimiel, J., & Munoz, A. (2011). Limitation of inverse probability of censoring weights in estimating survival in the presence of strong selection bias. *American Journal of Epidemiology*, ***173***(5).

Liang, K., & Zeger, S. (1986). Longitudinal data analysis using generalized linear models. *Biometrics*, ***73***(1), 13–22.

Molenberghs, G., & Verbeke, G. (2005). *Models for discrete longitudinal data* (2nd ed.). Springer Verlag.

Morden, J., Lambert, P., Latmer, N., Abrams, K. R., & Wailoo, A. (2011). Assessing methods for dealing with treatment switching in randomised controlled trials: a simulation study. *BMC Medical Research Methodology*, ***11***(4), 1186–1471.

Nielsen, J., Andersen, H., & Thomsen, P. (1998). Heart failure and echocardiographic changes during long-term follow-up of patients with sick sinus syndrome randomized to single chamber atrial or ventricular pacing. *Circulation*, ***97***, 987–995.

Rimawi, M., & Hilsenbeck, S. (2012). Making sense of clinical trial data: Is inverse probability of censoring weighted analysis the answer to crossover bias? *PubMed*, ***9***(3), 133–134.

Spruance, S., Reid, J., Grace, M., & M., S. (2004). Hazard ratio in clinical trials. *American Journal of Epidemiology*, ***48***(8), 2787–2792.

Sweeney, M. O., Hellkamp, A. S., Ellenbogen, & A., K. (2003). Adverse effect of ventricular pacing on heart failure and atrial fibrillation among patients with normal baseline qrs duration in a clinical trial of pacemaker therapy for sinus node dysfunction. *Circulation*, ***107***(23), 2932–2937.

Westreich, D., Cole, S. D., Tien, P., & et al. (2010). Time scale and adjusted survival curves for marginal structural cox models. *American Journal of Epidemiology*, ***171***(6), 550–560.

White, I. (2012). Methods for handling treatment switching: rank-preserving structural nested failure time models, inverseprobability-of-censoring weighting, and marginal structural models. *MRC Biostatistics Unit, Cambridge, UK*.

Zeger, S., & Liang, K. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*, ***42***(1), 121–130.

CHAPTER 6

APPENDIX

Table 6.1: *GEE parameter estimates (Empirical standard errors): Model (randomized treatment)*

Parameter	**Estimate (S.E)**	**95% Confidence Limits**	**Z**	**Pr**$>|Z|$
INTERCEPT	0.173(0.920)	[-3.542 ; 0.064]	-1.890	0.059
VISIT	-0.003(0.001)	[-0.006 ; -0.001]	-2.580	0.010
AGE	0.010(0.011)	[-0.011 ; 0.032]	0.930	0.354
GEN(Female)	-0.439(0.310)	[-1.047 ; 0.169]	-1.420	0.157
TRT_RAND (0)	-0.887(0.522)	[-1.909 ; 0.135]	-1.700	0.089
VISIT*TRT_RAND(0)	0.001(0.002)	[-0.002 ; 0.004]	0.580	0.564

Table 6.2: *GEE parameter estimates (Empirical standard errors): Model (treatment history)*

Parameter	**Estimate (S.E)**	**95% Confidence Limits**	**Z**	**Pr**$>|Z|$
INTERCEPT	-1.720(0.916)	[-3.515 ; 0.074]	-1.88	0.060
VISIT	-0.004(0.001)	[-0.006 ; -0.001]	-2.76	0.006
AGE	0.011(0.011)	[-0.010 ; 0.033]	1.02	0.308
GEN	-0.470(0.309)	[-1.074 ; 0.135]	-1.52	0.128
TRT	-1.119(0.509)	[-2.116 ; -0.123]	-2.20	0.028
VISIT*TRT	0.002(0.002)	[-0.001 ; 0.005]	1.08	0.280

Table 6.3: *Pooled logistic regression model for switching probability*

Effect	Wald ChiSq	PrÂ >Â ChiSq
VISIT	313.541	<.0001
PR_TRT	917.823	<.0001
GEN	44.116	<.0001
AGE	3.252	0.071
VISIT*PR_TRT	22.519	<.0001

Table 6.4: *Weighted Cox PH model: Parameter estimates (Empirical standard errors): Model (received treatment)*

Parameter	**Estimate (S.E)**	**95% Confidence Limits**	**Z**	**Pr**$>|Z|$
INTERCEPT	-1.634(0.933)	[-3.462 ; 0.194]	-1.750	0.080
VISIT	-0.004(0.001)	[-0.007 ; -0.001]	-2.870	0.004
AGE	0.011(0.011)	[-0.011 ; 0.034]	0.980	0.325
GEN	-0.505(0.322)	[-1.136 ; 0.127]	-1.570	0.117
TRT	-1.184(0.512)	[-2.188 ; -0.180]	-2.310	0.021
VISIT*TRT	0.002(0.002)	[-0.001 ; 0.005]	1.220	0.224

Auteursrechtelijke overeenkomst

Ik/wij verlenen het wereldwijde auteursrecht voor de ingediende eindverhandeling:
Estimate of treatment effect when the treatment changes in time.

Richting: **Master of Statistics-Biostatistics**
Jaar: **2014**

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,

Rukundo, Gilbert

Datum: **21/10/2013**