University of Padua
Department of Statistical Sciences
Master's degree in
Statistical Sciences

Efficient estimation of the marginal mean of recurrent events in randomized clinical trials

Supervisor: Prof.ssa Giuliana Cortese
Department of Statistical Sciences, University of Padua

Co-supervisor: Prof. Thomas Scheike
Section of Biostatistics, University of Copenaghen

Graduand Luca Genetti
Serial number 2021844

Academic year 2022/2023

Contents

Introduction

Recurrent events data are often encountered in biomedical settings, where individuals may also experience a terminal event such as death. Frequently, counting the observed number of adverse recurrent events is particularly relevant when one would like to assess the efficacy of a new treatment in reducing such events. In regard to this, a motivating data example is the LEADER study where randomized clinical trials were performed to assess whether the use of a drug based on Liraglutide reduces the number of myocardial infarctions (Marso et al., 2016). To address such a question a useful estimand for recurrent events data is the marginal mean of the cumulative number of recurrent events of a specific type, considering also the possible presence of a terminal event, as described in Cook and Lawless (1997); Ghosh and Lin (2000). A broader discussion of the analysis of recurrent events can be found in Cook et al. (2009); Cook and Lawless (2002). Recently, it was shown how the marginal mean number of recurrent events can be estimated efficiently by augmented estimators, which are then associated with more precise confidence intervals (Cortese and Scheike, 2022).

Recently, in the pharmaceutical industry there has been a renewed interest in studying recurrent events data more efficiently and correctly (Fritsch et al., 2021; Akacha et al., 2018; Schmidli et al., 2021). In regard to this, this thesis focuses on randomized clinical trials (RCTs), where it is often of interest to estimate the treatment effect of a new drug. In the RCT setting, the average causal treatment effect can be measured by the difference in the marginal means of the number of recurrent events between treated and untreated subjects. This difference

is an important estimand that can be estimated by comparing the simple marginal means estimator of each group, separately.

In this work, we show how to improve the efficiency of such estimand and present a novel doubly augmented estimator, which is enriched by two augmentation terms. The first augmentation term is related to missing data due to right censoring, and for each subject, uses history information that is collected dynamically over time and the auxiliary covariates. The second augmentation term exploits the randomization of the treatment. We demonstrate the important result that the two different augmentation terms are orthogonal, and thus contribute with different sources to reducing the variance of our estimators. As a consequence, we can study them separately. The two augmentation terms are based on working regression models and we show that, even if these models are misspecified, we still obtain a reduced variance for our doubly augmented estimators, at least asymptotically, compared to the simple non-augmented RCT estimator. In addition, the proposed estimators are consistent as long as the censoring adjustment is correct; see also Rosenblum and Steingrimsson (2016); Blanche et al. (2022).

We start by considering two different types of augmented estimators for the marginal mean of recurrent events: the augmented IPCW estimator that has been proposed by Cortese and Scheike (2022) in a more general context with respect to RCTs; an RCT augmented estimator based on semiparametric efficiency theory (Tsiatis, 2006), where randomization is the cause of the missing data setting. Then, to construct the proposed doubly augmented estimator, we take the augmented IPCW estimator and add the augmentation term of the RCT estimator that can improve the efficiency of the estimates furthermore. The first augmentation that improves the efficiency with respect to the right censoring, is based on a dynamic regression augmentation that is easy to compute, but as pointed out in Cortese and Scheike (2022), to get an improvement it is crucial that this is done dynamically over time. We provide a specific formula for the optimal gain that can be achieved

from this augmentation and demonstrate that the regression augmentation will always improve the asymptotic variance of the estimator. The second augmentation is achieved by utilizing the randomization and has a structure that is well understood, see for example (Tsiatis, 2006; Van der Laan and Robins, 2003; Robins and Rotnitzky, 1992). We also provide a specific formula for the optimal gain that can be obtained due to randomization.

The optimal gain in efficiency can only be obtained if in the augmentation term, the conditional model of the response, given covariates and treatment, is known. In practice, however, this is not feasible and we must use a working model. We show that if the working model is chosen appropriately, even if misspecified, we still obtain consistent estimators that will improve on the asymtotic variance, with respect to the simple RCT estimator. When the treatment difference is of interest then the working models should be chosen specifically for this purpose.

In particular, we propose two alternatives based on G-estimation: the first is a simple IPCW estimator and the second one comes from Gosh and Lin model with IPCW. We will apply these estimators to the marginal mean of recurrent events for treated, untreated groups and the difference of them. We will compare the efficiency between the simple marginal estimator without any augmentations, the one with one augmentation and the last one with the double augmentation, showing that we can gain in terms of efficiency, if the working model used for the augmentation is correctly specified and thanks to the orthogonality between the two augmented terms.

We will apply our methods to the LEADER clinical trial, a study that investigates the efficacy of a new treatment based on the Liraglutide drug in reducing cardiovascular disease in type-two diabetes patients. A randomized clinical trial was performed in which patients with type-two diabetes and high cardiovascular risk were randomly assigned to receive Liraglutide or placebo. Although the primary goal was to study the treatment effect on cardiovascular mortality, other cardiovascular

events of interest that can occur more than once over time are very relevant to study disease progression. Thus, it is also of great interest to investigate whether the treatment can reduce the number of cardiovascular recurrent events. See Furberg et al. (2022) for a discussion on different relevant estimands and methodology for this study, and Marso et al. (2016) for additional details. The LEADER trial considers two different outcomes of interest: the number of non-fatal myocardial infarctions (MI outcome), which is a classical recurrent events outcome, with the possible presence of death as a terminal event; the number of both non-fatal myocardial infarctions, non-fatal strokes and cardiovascular deaths (3-p MACE outcome), still in presence of death from other causes as a terminal event. Here, the 3-p MACE outcome is a so-called composite outcome that combines together the recurrent events and a type of terminal event of specific interest, see Mao and Lin (2016) for this approach. Consequently, in this context, non-cardiovascular death plays the role of a competing-risk event. The proposed estimation procedure for improving efficiency can also be applied to the composite outcome. We discuss how to treat this specific setting and show that it does not change so much from the arguments presented for the simple outcome.

Extensive finite sample simulation studies demonstrate that there are indeed important gains in efficiency in settings that mimic those of the LEADER data mostly for the composite 3-p MACE outcome that we considered in different simulations, since this type of events is more of interest respect the simple MI and the results we obtained were comparable.

The thesis is structured as follows. Chapter 1 describes the classical non-augmented estimator, the IPCW augmented estimator and the RCT augmented estimator, the doubly augmented estimator together with the possible choices for the working model of the augmentation terms. Chapter 2 shows the results of the simulation studies, for the 3-p MACE outcome; the performance of the proposed estimators in finite samples,

and in particular their gain in efficiency, are reported. The worked example about the LEADER clinical trial is described Chapter 3. Finally, Chapter 4 reports a conclusion with final remarks.

Model formulation

Let D denote the survival time (the terminal event), and let $N^*(t)$ count the number of recurrent events observed over a time-period $[0, t]$, where $t \leq \tau$. Due to the terminal event, we only observe the recurrent event processes up to $t \wedge D$, where $a \wedge b = \min(a, b)$, such that $N^*(t) = N^*(t \wedge D)$ because subjects will only have events while still alive. Observations may also be censored, thus only making it possible to observe the processes up to the censoring time C. Let us define $\Delta = I(D \leq C), T = D \wedge C$, and let $N(t) = N^*(t \wedge T)$ be the number of events observed while under risk and define the at-risk process $Y(t) = I(T \geq t)$. We also define as $G_c(t)$ the survival distribution of the censoring time C and for simplicity of notation, we assume that it does not depend on any of the covariates. In addition, let us consider the cumulative rate $\Lambda_c(t)$ and the counting process $N^c(t) = I(T \leq t, \Delta = 0)$ for the censoring time and the associated martingale given as $M^c(t) = N^c(t) - \int_0^t Y(s) d\Lambda_c(s)$.

We make the standard assumption that the right censoring is independent of D and $N^*(t)$. Then we consider the dichotomous treatment variable A that is equal to 1 for subjects treated with the new drug, and equal to 0 for those treated with placebo. In addition, we consider a vector of auxiliary covariates X. Let us assume that A and X are independent, and this is achieved due to randomization. Hence, we have that $P(A = 1|X) = P(A = 1) = \delta$. Then, in the causal inference framework, to define average causal treatment effects that will be discussed later, we introduce the potential outcomes notation for untreated and treated individuals (with assignment indicator $A = a$

with $a = 0, 1$) and assume no unmeasured confounding. The potential outcomes $N_i^*(t, a)$ and $Y_i(t, a)$ of an individual i, refer to the number of recurrent events and to the at-risk indicator that would be observed if the individual i was assigned to treatment a, for $a = 0, 1$.

Finally, we assume that our observations $(N_i(t), T_i, \Delta_i, A_i, X_i)$ are independent replicates of $(N(t), T, \Delta, A, X)$ for $i = 1, ..., n$ and are observed over the time interval $[0, \tau]$ where $\tau > t$. In addition, let $N_\bullet(t) = \sum_i^n N_i(t)$ and $Y_\bullet(t) = \sum_i^n Y_i(t)$ be the summary processes over the considered sample.

Our final scope is to estimate the marginal means of the number of recurrent events for treated and untreated, and the average causal treatment effect. For causal inference, we need to consider the expected mean $\mu(t, a) = E(N^*(t, a))$ of the potential outcome, interpreted as the marginal mean of the number of recurrent events in the population if treatment a was assigned to all the population. However, the randomization in the RCT context allows us to assume that the treatment assignment is independent of $N^*(t), D$ and X. Consequently, it holds that $\mu(t, a) = E(N^*(t, a)) = E(N^*(t)|A = a)$ and our key quantity of interest can be interpreted as the marginal mean of the number of recurrent events among those individuals treated with a. Moreover, under the independence assumption, the average causal treatment effect can be measured by the difference of the marginal means in the treated and untreated populations, as follows

$$\mu_{Diff}(t) = \mu(t, 1) - \mu(t, 0) = E(N^*(t)|A = 1) - E(N^*(t)|A = 0).$$

In the following, first we describe the simple (non augmented) estimator for the marginal mean number of recurrent events. Then, we illustrate and compare the two extended efficient estimators, the first one based on IPCW augmentation and the second one based on RCT augmentation. For ease of reading, the presentation of the simple estimator and IPCW augmented estimator is given, first, in general form with standard notation, and then it is extended and adapted to deal with the RCTs setting.

2.1 The Simple estimator for the marginal mean

The marginal mean number of recurrent events can be written as

$$\mu(t) = E(N^*(t)) = \int_0^t S(s)dR(s) \tag{2.1}$$

where $S(t) = P(D > t)$ is the probability of survival at time t and $dR(t) = E(dN^*(t)|D > t)$ is the recurrent events rate among survivors. This can correctly be considered a marginal mean because the increment given by $dR(t)$ at time t is not zero only if the terminal event time is such that $D > t$, otherwise it is equal to zero.

A simple nonparametric estimator for the marginal mean number of recurrent events given in (2.1) was proposed by Cook and Lawless (1997) and after developed by Ghosh and Lin (2000). Later on, this estimator will be regarded as a benchmark for comparing the augmented estimators with. The simple estimator has the form

$$\hat{\mu}(t) = \int_0^t \hat{S}(s)d\hat{R}(s) \tag{2.2}$$

where $\hat{S}(t)$ is the Kaplan Meier estimator and with $\hat{R}(t) = \int_0^t Y_\bullet^{-1}(s)dN_\bullet(s)$. It is asymptotically normal such that $n^{1/2}\{\hat{\mu}(t) - \mu(t)\}$ converges weakly to a mean-zero Gaussian process with an asymptotic variance that can be written as

$$\begin{aligned} &E\left(N^*(t) - \mu(t)\right)^2 \\ &+ \int_0^t E\left(\left[H(s,t) - E\left(H(s,t)|D > s\right)\right]^2 |D > s\right) S(s)\frac{\lambda_c(s)}{G_c(s)}ds \end{aligned} \tag{2.3}$$

where $H(s,t) = \int_s^t I(D > u)dN^*(u)$ is the observed number of recurrent events from s to $t \wedge D$. This variance can be consistently estimated by an IPCW estimator; see Cortese and Scheike (2022) for further details.

In the RCT setting, the quantities of interest are $\mu(t, a)$, for $a = 0, 1$, and they can be estimated by a simple RCT estimator that has the same form as $\hat{\mu}(t)$ given in (2.2), but computed only on observations of treatment group a as follows:

$$\hat{\mu}(t, a) = \int_0^t \hat{S}(s, a)\, d\hat{R}(s, a), \tag{2.4}$$

where $\hat{S}(s, a)$ is the Kaplan Meier estimator computed on the subsample with treatment a, and $\hat{R}(s, a) = \int_0^t Y_\bullet^{-1}(s, a) dN_\bullet(s, a)$ with $Y_\bullet(t, a) = \sum_i A_i Y_i(t, a)$ and $N_\bullet(t, a) = \sum_i A_i N_i(t, a)$ and $N_i(t, a) = N^*(t \wedge T, a)$. The estimator $\hat{\mu}(t, a)$ is asymptotically normal with variance given by equation (2.3), where in this case the involved quantities refer to the subpopulation with treatment a and the conditional expectations are written, given $D > t$ and $A = a$.

Consequently, a simple RCT estimator for the average causal treatment effect is obtained as $\hat{\mu}_{Diff}(t) = \hat{\mu}(t, 1) - \hat{\mu}(t, 0)$, i.e., the difference between the simple RCT estimators for the two treatment groups.

2.2 Introduction to inverse probability weighting

2.2.1 Inverse probability of treatment weighting

To describe the use of inverse probability censoring weighting we start from the inverse probability weight for treatment, that was borne in the observational studies, specially case-control ones. In such type of study, the goal is to see the difference of an outcome of interest between Treated (cases) and Untreated (controls) subjects. It can happen often that the number of case and control subjects are unbalanced; therefore, the results that we would obtain should be biased. To avoid this issue we can use a different technique to obtain a balanced sample of cases and controls, like matching 1:n. However, this method requires taking exactly n controls for every case and this should bring to a loss of information because we have to consider exactly $n \cdot n_{cases}$, and they could be lower than the number of controls in the original sample. Inverse probability of treatment weighting can bypass this problem. It proposes to use the entire number of Treated and Untreated but weighting their outcome in a way that is the inverse of their propensity score for the treatment's assignment. To explain this point, we consider Y to be the outcome of interest, A the dichotomous Treatment (A=1 for Treated and A=0 for Untreated) and X a vector of different covariates. Moreover, $Y_{A=1}$ and $Y_{A=0}$ are the outcome referred to the Treated and Untreated subjects respectively. Then we define the inverse probability weight, respectively, for Treated and Untreated as follow:

$$\frac{1}{P(A=1|X)}, \qquad \frac{1}{P(A=0|X)}. \tag{2.5}$$

The probabilities $P(A=1|X)$ and $P(A=0|X)$ can be estimated in a naive way by the proportion of Treated and Untreated subjects in

the sample, ignoring the covariates. Alternatively, a regression model for the binary variable A with covariate vector X can be fitted and used to predict the two propensity scores. Weighting the outcome Y as shown previously, we obtain two new outcomes as follows:

$$\tilde{Y}_{A=1} = \frac{1}{P(A=1|X)} Y_{A=1} \tag{2.6}$$

$$\tilde{Y}_{A=0} = \frac{1}{P(A=0|X)} Y_{A=0} \tag{2.7}$$

Using this technique, the new outcome is comparable to the one obtained from a sample with a balanced number of case and control subjects. However, we have to consider that the weights depend on the observed covariates X that could be also confounders, but they are also susceptible to unobserved confounders.

2.2.2 The use of inverse probability weighting in RCT setting

In this subsection we move to the context of randomized clinical trials. This setting is completely different from observational studies, where we can not know the experimental design of the study and the way with which the subjects have been selected. Instead, in a randomized clinical trial we decide the experimental design, and thus the way to select subjects and to assign the Treatment of interest. Therefore, in that RCT context we do not need a propensity score approach because randomization of subjects, that will be Treated and Untreated, avoids the problem of confounding and sample balance. However, often when we measure our outcome along a specific follow-up time like in time to-event analysis, or more specifically in recurrent events analysis, subjects can move out from the study before the end: this is the case of right censored data. In such a situation, using just observed outcomes, we could obtain biased results due to the censoring that precludes the observation

of the entire outcome. One approach that permits to solve this problem is based on an idea similar to the inverse probability of treatment weighting that has been explained above. It consists on considering the subjects that are not censored at a specific time t and weighting their outcome $Y(t)$ with the inverse probability of being uncensored at t, $G = P(C > t)$. Therefore, the new weighted outcome becomes:

$$\tilde{Y} = \frac{\Delta(t)Y(t)}{G(t)}, \tag{2.8}$$

where $\Delta(t) = I(C > t)$. This weighted outcome provides a new pseudo-population adjusted for the censoring. This consents to avoid biased results due to selection bias when we consider the simple observed outcome $Y(t)$ in presence of censoring. In general, we do not know the distribution of the censoring time and thus, the survival function $G(t)$ needs to be estimated. In this work we use the Kaplan-Meier non-parametric estimator to estimate $G(t)$.

2.3 IPCW estimator

Using the approach proposed firstly by Tsiatis (2006) that treated censored observations as missing data, we can rewrite the estimator given above in (2.2) as follows:

$$\begin{aligned}
\hat{\mu}(t,a) &= \int_0^t \hat{S}(s,a)\hat{d}R(s,a) \\
&= \sum_{i=1}^{n} \int_0^t \hat{S}(s,a)Y_i(s,a)\frac{1}{\sum_{i=1}^{n} Y_i(s,a)}dN_i(s,a) \\
&= \frac{1}{n}\sum_{i=1}^{n} \int_0^t Y_i(s,a)\frac{1}{\sum_{i=1}^{n} \hat{G}_c(s,a)}dN_i(s,a) \\
&= \frac{1}{n}\sum_{i=1}^{n} \int_0^t r_i(s,a)I(D_i \geq s,a)dN_i(s,a), \qquad (2.9)
\end{aligned}$$

where $G_c(t,a)$ is the survival distribution of censoring time C and $\hat{G}_c(t,a)$ the respective Kaplan-Meier estimation. This is the so-called IPCW estimator because it uses the inverse of survival probability of censoring time C to weight the original data, giving us an alternative estimator unbiased that takes in consideration the censoring time present in real data. Now we define the martingale process for the censoring time C as $M_i^C(t,a) = N_i^C(t,a)\int_0^t Y_i(s,a)d\Lambda_i^C(s,a)$, with $N_i^C(t,a) = I(T_i \leq t, \delta = 0)$. Using its property we can show that the corresponding normalised estimator $n^{1/2}\{\hat{\mu}(t,a) - \mu(t,a)\}$ can be written in the following form:

$$
\begin{aligned}
&n^{1/2}\{\hat{\mu}(t,a) - \mu(t,a)\} \\
&= n^{-1/2}(\sum_i \int_0^t \frac{Y_i(s,a)}{\hat{G}_c(s,a)dN_i(s,a)} - \mu(t,a)) \\
&= n^{-1/2}(\sum_i \int_0^t I(D_i > s)dN_i^*(s,a) - \mu(t,a)) \\
&- n^{-1/2}\sum_i \int_0^t [H_i(s,t,a) - E(H(s,t,a))]\frac{1}{G_c(s,a)}dM_i^C(s,a) + o_p(1)
\end{aligned}
\tag{2.10}
$$

where $E[H(s,t,a)] = (E[H_i(s,t,a)]I(D_i \geq s,a))/S(s,a)$ with $Hi(s,t) = \int_0^t I(D_i > u,a)dN_i^*(u)$.

Consequently, the variance of such estimator is equal to

$$
\begin{aligned}
&E(\int_0^t I(D_i > s,a)dN_i^*(s,a) - \mu(t,a))^2 + \\
&E(\int_0^t [H_i(s,t,a) - E(H(s,t,a))]\frac{1}{G_c(s,a)}dM_i^C(s,a))^2,
\end{aligned}
\tag{2.11}
$$

since the two terms are independent. The contribution of the Kaplan-Meier estimator for $G_c(s,a)$ to the influence function is the extra term involving $E[H(s,t,a)]$ in the martingale integral. As a consequence of this, the terms $H_i(s,t,a)$ are centred with respect to their conditional mean $E[H(s,t,a)]$, and thus the variance of the normalised estimator is reduced, if compared with the variance of the same estimator where the true $G_c(s,a)$ is used. See Cortese and Scheike (2022) for major details.

2.4 The IPCW augmented estimator

We showed above that the simple estimator of the marginal mean number of recurrent events can be rewritten as an IPCW estimator, which treats right censoring as a missing data problem. Using the results of Tsiatis (2006); Van der Laan and Robins (2003); Robins and Rotnitzky (1992), Cortese and Scheike (2022) proposed an extension of the simple IPCW estimator that consents us to improve the efficiency, using the theory of missing data developed in the context of semi-parametric theory. The efficient IPCW estimator is based on an augmented term and is given as:

$$\begin{aligned}\tilde{\mu}(t) &= \hat{\mu}(t) + \frac{1}{n}\sum_i \int_0^t \frac{L_i^{eff}(s,t)}{\hat{G}_c(s)} d\hat{M}_i^C(s) \\ &= \hat{\mu}(t) + \frac{1}{n}\sum_i \int_0^t \frac{L_i^{eff}(s,t)}{\hat{G}_c(s)} dN_i^C(s), \end{aligned} \tag{2.12}$$

where the second term, the so-called augmentation term, is responsible for the improved efficiency of the simple estimator $\hat{\mu}(t)$. The most efficient estimator is obtained when
$L_i^{eff}(s,t) = E[H_i(s,t)|\mathcal{H}_i(s), D_i > s)] - E(H(s,t)|D > s)$,
with $\mathcal{H}_i(s)$ being the history of the ith subject observed up to time s.

In Appendix A.1, we show that the efficient IPCW estimator reduces the variance of the simple RCT estimator by the following quantity:

$$\int_0^t E[\left(E\left(H_i(s,t)|\mathcal{H}_i(s), D_i > s\right) - E\left(H_i(s,t)|D_i > s\right)\right)^2 |D_i > s] \, S(s) \frac{\lambda_c(s)}{G_c(s)} ds. \tag{2.13}$$

Consequently, this reduction depends on the variation of the conditional mean given the history.

From theoretical results, we know that the estimator $\tilde{\mu}(t)$ with $L_i^{eff}(s,t)$ defined as above, is optimal in terms of efficiency, in the sense that its asymptotic variance is the smallest one among the class of regular asymptotically linear estimators. However, in practice, it will

often be impossible to obtain the full efficiency gain since the conditional mean $L_i^{eff}(s,t)$ is computationally difficult to obtain. Therefore, we here use the suggestion in Bang and Tsiatis (2000) and extend it to dynamic regression with time-varying coefficients. Then, to approximate $L^{eff}(s,t)$, we consider dynamic regression modelling of $H(s,t)$ based on J predictors $e^T(s,t) = (e^1(s,t), ..., e^J(s,t))$. These predictors will be auxiliary covariates and internal time-dependent covariates such as, for example, the number of recurrent events up to time s, and possible interactions between these two types of covariates. This regression model is clearly an approximation and will often provide a substantial gain in efficiency, but it can not reach the efficiency bound of the most efficient estimator. The augmented IPCW estimator based on dynamic regression is computed as

$$\tilde{\mu}_r(t) = \hat{\mu}(t) + \frac{1}{n}\sum_{i=1}^{n}\int_0^t \frac{\gamma(s,t)^T(e_i(s) - \bar{e}(s)}{\hat{G}_c(s)} dN_i^C(s), \qquad (2.14)$$

where $\gamma(s,t)$ is a J-dimensional vector of time-varying regression coefficients, and $\bar{e}(s) = \sum Y_i(s)e_i(s)/Y_\bullet(s)$ is the at-risk average of the predictors $e_i^T(s) = (e_i^1(s,t), ..., e_i^J(s,t))$. This estimator, although not optimal, will be more efficient than the simple RCT estimator and, when auxiliary covariates as well as heterogeneity from the history, such as the number of recurrent events observed at time s, are chosen opportunely as predictors, the gain in efficiency may be substantially high.

Our scope is to minimize the variance of the normalized estimator $n^{1/2}(\tilde{\mu}_r(t) - \mu(t))$. Its influence function can be written appropriately and the variance is minimized when the estimator of $\gamma(s,t)$ is obtained by regressing $H_i(s,t) - E(H(s,t)|D > s)$ on $e_i(s) - \bar{e}(s)$. An explicit formula of this optimal estimator and further theoretical details are given in Cortese and Scheike (2022). In addition, the variance of $\tilde{\mu}_r(t)$ can be estimated consistently by

$$\widehat{var}(\tilde{\mu}_r(t)) = \widehat{var}(\hat{\mu}(t)) - n^{-1}\int_0^t \hat{\gamma}(s,t)^T\tilde{\Sigma}(s)\hat{\gamma}(s,t)\frac{1}{\hat{G}_c^2(s)Y_\bullet(s)} dN_\bullet{}^c(s), \qquad (2.15)$$

where $\widehat{var}(\hat{\mu}(t))$ is the estimated variance of the simple estimator of section 2.1 without the augmentation term and $\tilde{\Sigma}(s)$ is the sample variance of the $e_i(s)$, for $i = 1, \ldots, n$, among survivors. As a consequence, we note that the dynamic-regression augmentation will always improve the asymptotic variance, which is reduced by the second term in (2.15). This term is indeed the estimator of the variance reduction given in (2.13).

In the RCT setting, the augmented IPCW estimators described in this subsection can be considered separately for the two treatment groups. In this case, all the involved expected values are taken conditionally on $D > t$ and $A = a$. The resulting estimators $\tilde{\mu}_r(t, a)$, for $a = 0, 1$, are computed as given in (2.14), by using only observations who received treatment a. The regression estimators, here denoted with $\hat{\gamma}(s, t, a)$, are computed doing regression with only observations in group a. The estimators $\tilde{\mu}_r(t, a)$ are asymptotically normal and similarly, the variance estimator in (2.15) is computed separately for each treatment group. Finally, an efficient IPCW estimator for the average causal treatment effect can be obtained by the simple difference between the estimators for the two treatment groups, i.e., by $\tilde{\mu}_{r,Diff}(t) = \tilde{\mu}_r(t, 1) - \tilde{\mu}_r(t, 0)$.

2.5 The RCT Augmented estimator

We now consider the RCT setting and here review some results about efficient estimators in this case. As already pointed out, due to the orthogonality that we will show in the next section, we can discuss RCT estimators in the case of fully observed data disregarding the right censoring. Doing so, the response of interest is the fully observed number of recurrent events up to time t, denoted here with $Z_t = N^*(t)$. In the specific setting of randomized clinical trials, we can gain important efficiency by using the fact that we have randomization, that is, that A and X are independent.

Our interest is devoted to estimating the marginal mean number of recurrent events among the treated subjects, $\mu(t,1) = E(Z_t|A=1)$, and among the untreated individuals, $\mu(t,0) = E(Z_t|A=0)$. Note that the response of interest and its marginal mean depend on t, but for the estimators presented in this subsection, t is considered fixed. In constructing an RCT estimator for the two marginal means, the efficiency gain from randomization is obtained by an RCT augmentation that involves the probability of treatment assignment δ, the observed treatment indicator A, and a working model necessary to estimate a conditional mean of the response of interest.

We here present the estimator for the marginal mean related to the treated arm, while the estimator for the untreated arm is obtained following similar arguments. We start by considering the inverse probability weighted complete-case estimator, $\hat{\mu}_c(t,1) = n^{-1}\sum_{i=1}^{n} A_i Z_{t,i}/\delta$. Then, to obtain the augmented version of this estimator, recall that the i-th efficient influence function is equal to:

$$\begin{aligned}\varphi(Z_{t,i}) &= \frac{A_i}{\delta}Z_{t,i} - \frac{A_i-\delta}{\delta}E(Z_{t,i}|X_i, A_i=1) - \mu(t,1)\\ &= \frac{A_i}{\delta}\left(Z_{t,i} - E(Z_{t,i}|X_i, A_i=1)\right) + E(Z_{t,i}|X_i, A_i=1) - \mu(t,1)\\ &= \frac{A_i}{\delta}\left(Z_{t,i} - \mu(t,1)\right) - \frac{A_i-\delta}{\delta}\left(E(Z_{t,i}|X_i, A_i=1) - \mu(t,1)\right);\end{aligned} \tag{2.16}$$

see Tsiatis (2006). Therefore, in absence of right censoring, the efficient estimator of $\mu(t,1)$ is

$$\tilde{\mu}_{rct}(t,1) = \hat{\mu}_c(t,1) - \frac{1}{n}\sum_i \frac{A_i - \delta}{\delta} E(Z_{t,i}|X_i, A_i = 1) \qquad (2.17)$$

This is an augmented IPW estimator and we note that the augmentation term depends on the marginal mean $E(Z|X, A = 1)$. The normalized estimator of (2.17), $n^{1/2}(\tilde{\mu}_{rct}(t,1) - \mu(t,1)) = n^{-1/2}\sum_i \varphi(Z_{t,i}) + o_p(1)$, is also asymptotically normal with variance given as

$$E[\frac{A}{\delta}(Z_t - \mu(t,1))]^2 - E[\frac{A-\delta}{\delta}(E(Z_t|X, A = 1) - \mu(t,1))]^2; \quad (2.18)$$

see Appendix A.2 for details. Therefore, the reduction in variance due to randomization is given by the last term of the previous expression, and is large if the variance of $E(Z_t|X, A = 1)$ is large. In practice, however, we must use a working model for $E(Z_t|X, A = 1)$ that we denote $h_t^*(X, A = 1)$. First we note that since we have randomization, the estimator will be unbiased, or consistent, no matter what working model that is used. This is direct consequence of the double-robustness property of the efficient influence function that will be unbiased solely due to the randomization. Using a working model we can compute the variance of the influence function that is proportional (in h_t^*) to

$$\left(\frac{A}{\delta}(E(Z_t|X, A = 1) - h_t^*(X, A = 1))^2\right) \propto E\left(\frac{A}{\delta}(Z_t - h_t^*(X, A = 1))^2\right).$$

As a consequence we have shown that indeed the efficient choice of $h_t^*(X)$ is the conditional mean, and that any working model will lead to variance reduction compared to the simple mean, if the working model is found such that it minimizes $E(\frac{A}{\delta}(Z_t - h_t^*(X, A = 1)))^2$. That is, under regularity conditions, when it solves the estimating equation $E(A \cdot Dh_\beta^*(\beta, X, A = 1)(1, X)^T(Z_t - h_t^*(\beta, X, A = 1)) = 0$. We stress that the working model contains a baseline parameter. Further, as a linear working model is found to solve $\sum_i A_i(Z_{t,i} - h_t^*(\hat{\beta}, X_i, A = 1)) = 0$ then $\hat{\rho}_1 = n^{-1}\sum_i h_t^*(\hat{\beta}, X_i, A = 1)$ is the G-estimator. In addition

when solving the score equation in practice, to deal with the censoring, we must use a IPCW adjusted response instead of the response Z_t.

We finally make a few remarks about the situation when interest is specifically on estimating the treatment effect, that is the difference in the marginal means between the treated and untreated subject, $\nu = \rho_1 - \rho_0$. All influence function for ν are given by

$$\frac{A}{\delta}Z_t - \frac{1-A}{1-\delta}Z_t - (A-\delta)h_t(X) - \nu \tag{2.19}$$

see Tsiatis (2006) page 130-131 and the efficient choice of $h_t(X)$ is $h_t^{eff}(X) = E(Z_t|X, A=1)/\delta + E(Z_t|X, A=0)/(1-\delta)$. Now computing the variance of the influence function based on the working model h_t^* we get that it is proportional to (see Appendix A.5):

$$\propto E\left(h_t^*(X) - h_t^{eff}(X)\right)^2 \propto E\left(h_t^*(X) - \tilde{Z}_t(A)\right)^2$$

where $\tilde{Z}_t(A) = AZ_t/(\delta^2) + Z_t(1-A)/(1-\delta)^2$. When $\delta = 1/2$ then $Z_t(A) = Z_t/\delta^2$. From this direct calculation we see that indeed the optimal choice of h_t is h_t^{eff} and we also note that we are able to guarantee to reduce variance when the working model solves the estimating equation $E((1,X)^T D_\beta h_t^*(\beta, X)(\tilde{Z}_t(A) - h_t^*(\beta, X))) = 0$.

2.6 Doubly augmented estimator

We now combine the censoring and the RCT augmentations of the two previous sections. A key result for doing this is that we establish that the two augmentations are orthogonal, that gives additive contributions to the variance.

Taking the starting point of the previous section, we start with an IPCW weighted response version of the fully observed response. Thus letting $Z_t^{ipcw} = \int_0^t Y(s)\frac{1}{G_c(s)}dN(s,1)$, we consider the influence function

$$\frac{A}{\delta}Z_t^{ipcw} - \frac{A-\delta}{\delta}E(Z_t|A=1,X) - \rho_1, \tag{2.20}$$

that can be made more efficient as follows by a censoring augmentation:

$$\begin{aligned}&\frac{A}{\delta}\left(Z_t^{ipcw} + \int_0^t \frac{E[H(s,t)|\mathcal{H}(s), D>s, A=1]}{G_c(s)}dM^C(s) - \rho_1\right)\\ &-\left(\frac{A-\delta}{\delta}E(Z_t|A=1,X) - \rho_1\right).\end{aligned} \tag{2.21}$$

Using the equality $Z_t^{ipcw} = Z_t - \int_0^t H(s,t)\frac{1}{G_c(s)}dM_c(s)$, we can rewrite the influence function as

$$\begin{aligned}&-\frac{A}{\delta}(\int_0^t (H(s,t) - E[H(s,t)|\mathcal{H}(s), D>s, A=1])\frac{1}{G_c(s)}dM_c(s))\\ &+\frac{A}{\delta}(Z_t-\rho_1) + \frac{A-\delta}{\delta}\left(E(Z_t|A=1,X)-\rho_1\right).\end{aligned} \tag{2.22}$$

A detailed derivation is provided in Appendix A.4. Due to the independent censoring assumption, the two terms subtracted in the two pairs of the round brackets in 2.21 are orthogonal and thus we can compute the variance of the influence function by the sum of the variance of the two terms. This is also an efficient influence function. We therefore propose to use the regression augmentation developed by Cortese and Scheike (2022) to reduce the variance due to the right-censoring term, and to use a working model to reduce the variance of the second term, thus combining what we saw in the two previous sections.

With these choices we thus solve the estimating equations and then obtain

$$\hat{\rho}_1 = \tilde{\mu}(t,1) - \frac{1}{n}\sum_i^n \frac{A_i - \delta}{\delta} h_t^*(X_i, A=1). \tag{2.23}$$

If the working model h_t^* is chosen as we argued in the previous section, we are guaranteed, at least asymptotically, to get a variance reduction from both the augmentations. We computed the variance of this estimator by utilizing again that the censoring augmentation in the first term is orthogonal to other terms in the expression. Therefore, we derived the variance of the estimator from its corresponding influence function in the following way:

$$\begin{aligned} &\hat{var}(\tilde{\mu}(t,1)) + E\left(\left(\frac{A-\delta}{\delta} h_t^*(X_i, A=1) - \rho_1\right)\right)^2 \\ &-2E\left(\frac{A(A-\delta)}{\delta^2}\left(Z_t^{ipcw} - \rho_1\right)\left(h_t^*(X_i, A=1) - \rho_1\right)\right), \end{aligned} \tag{2.24}$$

see Cortese and Scheike (2022) for detailed derivation of $\hat{var}(\tilde{\mu}(t,1))$.

Similarly, when aiming directly at estimating the treatment effect ν then IPCW augmenting all influence functions for ν we get

$$\begin{aligned} &\frac{A-\delta}{\delta(1-\delta)} Z_t - (A-\delta) h_t^*(X) - \nu \\ &+ \frac{A-\delta}{\delta(1-\delta)}\left(\int_0^t \frac{H_i(s,t) - E[H_i(s,t)|\mathcal{H}_i(s,a), D_i > s]}{G_c(s)} dM^C(s)\right) \end{aligned} \tag{2.25}$$

where again the optimal working model is

$$h_t^*(X,A) = E(Z_t|X, A=1)/\delta + E(Z_t|X, A=0)/(1-\delta)$$

. Importantly, the two augmentations are still orthogonal and we are sure to obtain a variance reduction if we choose the working models for the RCT augmentation and the censoring augmentation directly as the

solution h^* that minimizes $E((\tilde{Z}(A) - h^*(\beta, X))^2$. Thus we need to solve the equation

$$E(D_\beta h_t^*(\beta, X) \cdot (1, X)^T (\tilde{Z}_t(A) - h_t^*(\beta, X)) = 0.$$

When $\delta = 1/2$ then

$$\tilde{Z}_t(A) = AZ_t/(\delta^2) + Z_t(1 - A)/(1 - \delta)^2 = Z_t/\delta^2$$

and we will then in practice use Z_t^{ipcw} instead of Z_t. Combining the two doubly augmented estimators for the treatment and non-treatment arm we obtain the following estimator for the treatment effect:

$$\begin{aligned}\hat{\nu} = &\frac{1}{n}\left(n_1 \frac{1-\delta}{\delta(1-\delta)}\tilde{\mu}(t,1) - n_0 \frac{\delta}{\delta(1-\delta)}\tilde{\mu}(t,0)\right) \\ &- \frac{1}{n}\sum_{i=1}^{n}(A_i - \delta)(h_t^*(X_i, 1) - h_t^*(X_i, 0)).\end{aligned} \tag{2.26}$$

2.6.1 Working model's choice

As shown above, to construct the RCT augmentation we need a working model to approximate the conditional mean $E(Z_t|X, A = a)$ on which we can note that the treatment is fixed with a that can take the values $0, 1$. Therefore, we need to estimate each working model separately for the Treatment and the Untreatment groups (stratified model for the Treatment variable), according with the arguments presented in section 2.5. Nevertheless, we can also fit the working model jointly, considering at the same time both Treated and Untreated subjects. Obviously in this case we have to add the indicator variable and the respective coefficient referred to the Treatment. Following the first way we guaranteed theoretically and asymptotically that we can improve efficiency, but we adapted both the way to see in our practical context how the estimators differently perform. Here we consider two standard working models in the context of recurrent events. To estimate the parameters of such working models we need to use IPCW based estimating equations.

As a candidate working model h_t^*, we shall consider the simple marginal mean model

$$E(N(t)|X, A = a) = \exp\left(\beta_0 + \beta_X^T X\right),$$

that is a stratified model, or

$$E(N(t)|X, A) = \exp\left(\beta_0 + \beta_1^T A + \beta_X^T X\right),$$

that is a joint model with regression parameters $\beta_0, \beta_1, \beta_X$. To fit these models we use the IPCW adjusted response introduced in the previous sections, Z_t^{ipcw}, and then solve relevant estimating equations based on the suggested working model. We also note that if the estimator is constructed appropriately, using this working model, we are guaranteed to improve the variance (asymptotically) even if the working model is misspecified.

Another popular choice is the Gosh-Lin model (Ghosh and Lin, 2002) that assumes

$$E(N(t)|X, A = a) = \Lambda_0(t) \exp\left(\beta_X^T X\right),$$

that is a stratified model, or

$$E(N(t)|X, A) = \Lambda_0(t) \exp\left(\beta_1^T A + \beta_X^T X\right),$$

that is a joint model, for $t \in [0, \tau]$, $\Lambda_0(t)$ is a non-parametric baseline and β_1, β_X are the regression coefficients. Using one of these models as working model, and estimating its parameters with the standard IPCW estimators, we are not guaranteed that this choice will improve the variance of the estimator unless the working model is correctly specified. We illustrate with simulations that it may lead to biased estimates when the model does not hold over the entire time range. Further we are not guaranteed that there will be a variance reduction due to the RCT augmentation, even though in practice this will often be the case if the model fits well.

The two models have a different formulation of the baseline. IPCW estimator is a completely specified model, indeed also the baseline has a parametric form and it estimates the mean of recurrent event only for the specific time t and not along all the time frame. But the good property of this model states in its robustness in case of model misspecification, if the censoring distribution $G_c(s)$ is correctly estimated. On the other hand, Gosh-Lin IPCW model is based on a semi-parametric approach, since the baseline $\Lambda_0(t)$ is unspecified and so it is estimated in a non-parametric way. However, since it can vary on t, it consents to estimate the outcome throughout the whole time frame. Unfortunately, this model is not consistent in case of misspecification.

For this reason and since in this work we are focused on the marginal mean estimation at a specific time, we generally should prefer simple fixed time point IPCW working model but we will also apply the Gosh-Lin model to compare the two methods.

Importantly, we also note that since the randomization is carried out according to the design, then even when both models are correctly specified then the efficient influence function will lead to the same asymptotics irrespective of the working model being fitted by the Ghosh-Lin estimator or as the simple IPCW estimator. This is due the fact the fact that the influence function related to $\frac{A}{\delta}Y - \frac{A-\delta}{\delta}F(\hat{\beta}, X, A = 1) - \rho_1$, will lead to the same asymptotics as when β_0 for the working model is known.

Simulations

In this section we will apply the considered estimators to data that resemble the LEADER data. This will give some indication of the possible efficiency gain in such a setting. For all considered estimators we also study the performance of the estimated standard errors by, for example, computing the coverage probabilities of the constructed confidence intervals. As introduced at the beginning of this work, the process of interest, specially when we consider the 3p-MACE event, is characterized by a combination between a recurrent events process (non fatal strokes and myocardial infractions) and a terminal event (cardiovascular deaths). For this reason, we used two main approaches in the simulations. The first is to consider cardiovascular deaths as both a recurrent event and a terminal event, so it is incorporated in both the processes. Therefore, in this case we have only recurrent events, deaths and censorings. Instead, in the second approach we keep the recurrent events (non fatal stroke and myocardial infraction) and cardiovascular deaths separated. Now, the process of interest is the sum of a recurrent process and a specific terminal event, while the non cardiovascular death and the censoring remain unchanged. Despite the simulated results in term of marginal mean estimates and their efficiency are very similar between these settings, as we will show in the following, the second approach is closer to the real case and it's more correct, at least from a theoretical point of view.

3.1 Standard simulations

We first estimated the rate among survivors of the recurrent events using a Cox model that included treatment as well as a selected set of covariates that was known to be of importance for the risk of strokes. This lead to a baseline rate $k_1\lambda_r(t)$ and a set of regression coefficients β_r. The rate of the terminal event was also described with a Cox model with the same covariates, denoting the rate as $\lambda_d(t)$ and the set of regression coefficients β_d. We show the coefficients of the two Cox models related to a set of covariates selected based on subject matter knowledge, in Table 3.1. The covariates were Placebo (no Liraglutide), MIFL (Myocardial Infarction), KIDFL (Chronic Kidney Failure), REVASFL (Revascularization), AGE (Age), HBA1CBL (HbA1c at Baseline), LDL1BL (Calc. LDL Cholesterol at Baseline), HDL2BL (HDL Cholesterol at Baseline), EGFREPB (eGFR at Baseline), DIABPBL (Diastolic BP at Baseline), DIABDUR (Diabetes Duration), and SYSBPBL (Systolic BP at Baseline).

	Cox Model Death		Cox Model Recurrent	
	Estimate	Std.Err	Estimate	Std.Err
Placebo	0.29	0.09	0.21	0.06
sexMale	0.05	0.11	0.15	0.07
miflYes	0.71	0.11	0.56	0.07
kidflYes	-0.02	0.14	0.07	0.09
revasflYes	-0.09	0.11	0.20	0.07
age	0.03	0.01	0.02	0.00
hba1cbl	0.17	0.03	0.10	0.02
ldl1bl	0.19	0.05	0.17	0.03
hdl1bl	-0.42	0.17	-0.27	0.11
egfrepb	-0.02	0.00	-0.01	0.00
sysbpbl	0.00	0.00	0.01	0.00
diabpbl	0.01	0.01	-0.00	0.00
diabdur	0.01	0.01	0.01	0.00

Table 3.1: Regression coefficients for Cox models for death and recurrent events.

	Cox Model Death		
	Estimate	Std.Err	P-value
History	0.82	0.06	0.00
Placebo	-0.16	0.07	0.03
sexMale	-0.11	0.08	0.21
miflYes	0.43	0.09	0.00
kidflYes	-0.09	0.12	0.47
revasflYes	-0.11	0.09	0.22
age	0.04	0.01	0.00
hba1cbl	0.17	0.02	0.00
ldl1bl	0.13	0.04	0.00
hdl1bl	-0.17	0.14	0.21
egfrepb	-0.02	0.002	0.00
sysbpbl	-0.001	0.003	0.62
diabpbl	0.009	0.005	0.05
diabdur	0.002	0.004	0.73

Table 3.2: Regression coefficients for Cox model for death with historical number of recurrent events as a predictor.

The censoring times were simulated from a translated exponential distribution, as approximation of the censoring present in the data, see Figure 3.1 where the black curve is the censoring distribution in the data. We also considered a situation where censorings appeared earlier in time to see how the efficiency gain was improved in this case, the red curve of Figure 3.1. We varied k_1 parameter like $k_1 = 1, 2.5, 5, 10$ to increase the number of recurrent events ($k_1 = 1$ is referred to the baseline of the original LEADER data). Furthermore, to allow also dependence between recurrent and death processes we add a common frailty variable Z gamma distributed with variance equal to 1. Thus, in this setting the two baselines become $Zk_1\lambda_r(t)$ and $Z\lambda_d(t)$. To investigate the presence of the heterogeneity, we estimated a Cox model for death's risk using the history of recurrent events experienced by the subjects as predictor, also adjusting with the auxiliary covariates X. Since the regression's coefficients were significantly high, specially the one related to the historical number of recurrent events (see Table 3.2), we confirmed the

presence of dependence between the two processes. We now simulated data from these baselines given the covariates, that were resampled from the LEADER trial to get a realistic covariate distribution. For each setting, marginal mean and treatment effect have been estimated at two specific time: 1500 and 1700 days. For every estimand (marginal mean for untreated, treated arm and their difference), we show the following estimators: the censoring augmented estimators (that is the IPCW augmented estimators described in section 2.4), the two doubly augmented estimators with either IPCW or Gosh-Lin as working models, the G-estimators based on the two working models alone and the standard estimator (this one only for the tables about marginal means and estimated coverage). For such estimators, we fitted the working models in the two ways discussed in section 2.6.1: the stratified and the joint models. In all the settings, in the doubly augmented estimators we used both the history process and the covariates X as predictors to compute the censoring augmentation (model 2.14 in section 2.4). Every estimate (marginal mean, sample variance, estimated standard error's coverage) is based on 10.000 replications with sample size equal to the original LEADER data one, 9170. In the tables we let WM indicates the Working Model that can be simple IPCW or Gosh-Lin. $Time$ is the specific time at which the marginal means estimate is computed. k_1 is the multiplicative factor for $\lambda_r(t)$ that increases the number of recurrent events. Regarding the type of models, we refer to $Model0$ with the one that assumes independence between the two baselines $k_1\lambda_r(t)$ and $\lambda_d(t)$ given covariates. Instead, $Model1$ has additional dependence via the frailty variable Z.

In the following tables we will show the simulations results for the setting related to $Model1$, since this case is closer to the real one, as explained above. We will show the efficiency gain, coverage probability of the estimated standard errors and the marginal means for the two groups as well as the treatment effect (i.e. the difference between the groups). We expressed the efficiency gain as the ratio of the variance

of each estimator relative to that to the standard estimator.

We note in general that as expected all estimators provided gain compared to the standard estimator, specially when we are considering the standard censoring distribution, see Tables 3.3 and 3.4. Furthermore, the gain in efficiency increases when we scale up k_1, thus the number of recurrent events. However, this improvement is more evident when we fitted the working model jointly respect to fit it separately. Moreover, while adapting the stratified working models the doubly augmented estimator provides the best gain among all the estimators and the G-estimators (simple IPCW and Gosh-Lin) perform quite similar. Fitting the joint working models we can see that G-estimator Gosh-Lin reaches the highest gain among all the estimators, mostly increasing k_1. This can be due to the reason that the Gosh-Lin model is based on more restricted assumptions regarding the proportionality of covariates' effect mostly. Therefore, when this assumptions hold we can gain more in efficiency, also more than the doubly augmented estimators. On the other hand, it's not robust in case of assumptions' violations, as we will show in section 2.3. About censoring augmented estimators, we can note they do not improve efficiency substantially, even when scaling up k_1. Nevertheless, changing censoring distribution choosing the simple exponential (red curve of Figure 3.1), the results become quite different. First of all, looking at Tables 3.7 or 3.8 we can note note that the censoring augmented estimator performs better in this case, especially when scaling up k_1. This is reasonable since such augmentation exploits the censoring time to gain in term of efficiency. Indeed in this case, it starts at the beginning of the follow-up period, thus at the specific time considered (1500, 1700) we can collect several censored subjects and this increases the potential gain this type of estimators can give. Moving to the G-estimators, we note some similarities and differences between 3.7 and 3.8. About similarities, Gosh-Lin G-estimator performs quite better then the respective simple IPCW, and this behaviour is more present when we fitted the joint working model. However, specially in

this setting we note that simple IPCW G-estimators perform quite worse respect to the standard estimator, our benchmark. Indeed, the gain's loss increases as the number of recurrent events become higher, instead of Gosh-Lin model that provides efficiency also in this case. On the other hand, when we fitted the working models separately, both the two G-estimators reduce the variance respect to the standard one. Moreover, in this case the best gain is reached by the doubly augmented estimator and it's higher to the one obtained by the same estimator but using the joint working model for the rct augmentation. Therefore, this confirms the argument developed in chapter 2: fitting the working models separately we guarantee that we can reduce the estimates' variance, as we saw in all the simulations' settings tested. Despite this, we can gain in efficiency also fitting one working model jointly in several situations but not always, as we pointed out above.

About coverages and marginal mean estimators, we'll show the results only related to the setting with stratified working models, omitting the tables with the joint working models, since we obtained very similar values.

In Table 3.5 we show the observed coverage of 95% confidence intervals for every estimated standard error related to the respective estimator. We can see that the coverages are close to the nominal value 95% in almost all the settings, specially the ones related to the doubly augmented estimator. where it's up to 94%. To compute the coverage, since that we didn't know the true value, we needed to estimate it for each estimand. Therefore, we estimated each one as the sample mean of the marginal means obtained from 5000 iterations, each one with a larger sample size equal to 30.000.

In Table 3.6 we show the results for the marginal means for the two arms (Treated and Untreated subjects) and the treatment effect. We observed that all means are in agreement, and that they are all unbiased estimators. We recall that all the estimators except the Ghosh-Lin G-estimator should be unbiased by construction. We can see that the

estimates obtained with the G-estimators and the Doubly augmented estimators are very closed to the standard ones, there are some differences in the third/fourth decimal position. Therefore, we can conclude that the estimators proposed are substantially unbiased. Moreover, we have to point out that also Ghosh-Lin model seems to be unbiased, sign that probably in this setting (LEADER data) the assumptions on which it's based hold. About this, we will show in next section that this is not always valid. The same results are reported for $Model0$ in Appendix B.1. We can note in general that the doubly augmented estimators performs better in $Model0$ compared to $Model1$. Nevertheless, for the censoring augmented estimators we can observe an opposite behavior specially with the simple censoring distribution. Indeed, for such estimators, the heterogeneity provides an efficiency's increment slightly better when the number of recurrent event increases. That's a confirmation of the considerations done by Cortese and Scheike (2022).

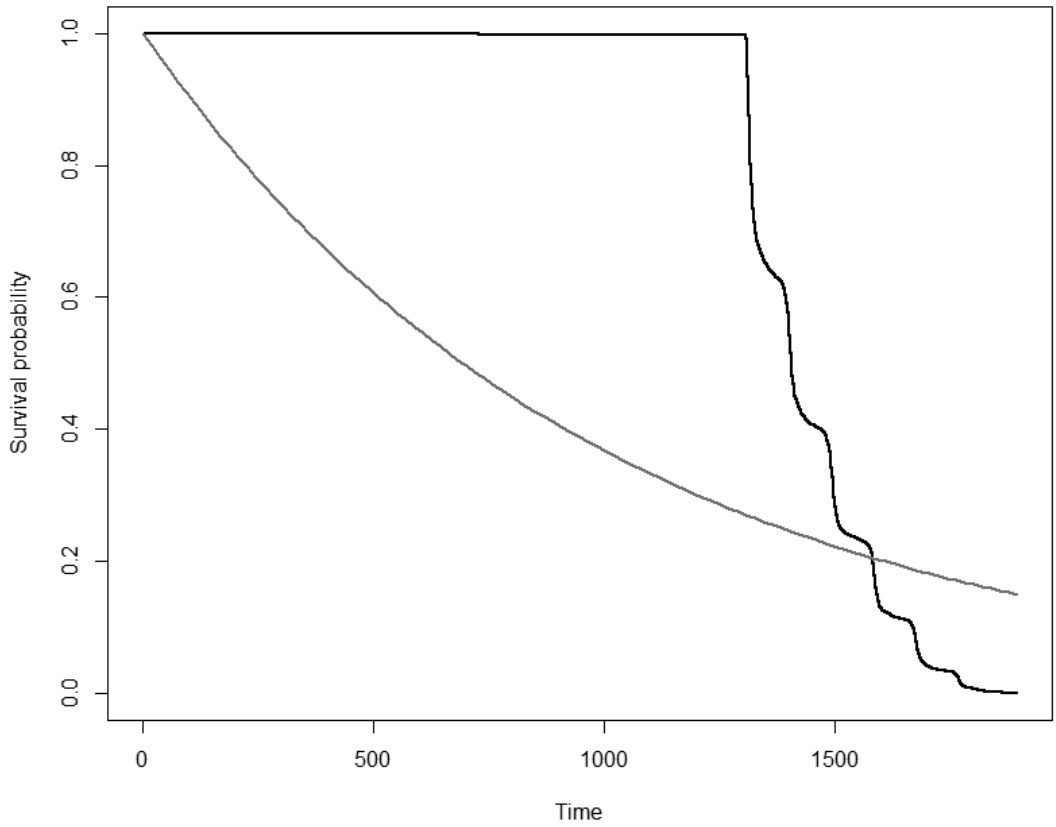

Figure 3.1: Survival distribution of censoring time of LEADER data (black curve) and simple exponential censoring distribution (red curve).

Table 3.3: Efficiency measure: ratio between sample variance of the main estimators and sample variance of standard estimators. Stratified working models for Treatment variable.

Model 1			Censoring aug. estimators			G-estimators			Doubly aug. estimators		
$Time$	k_1	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	1	IPCW	0.995	0.995	0.994	0.989	0.994	0.979	0.985	0.988	0.973
		GL	-	-	-	0.988	0.992	0.977	0.985	0.987	0.973
	2.5	IPCW	0.997	1	0.999	0.979	0.979	0.958	0.977	0.979	0.957
		GL	-	-	-	0.977	0.978	0.954	0.975	0.977	0.953
	5	IPCW	0.997	0.999	0.998	0.967	0.986	0.954	0.965	0.984	0.951
		GL	-	-	-	0.968	0.986	0.955	0.965	0.984	0.952
	10	IPCW	0.999	0.999	0.998	0.971	0.978	0.947	0.97	0.976	0.945
		GL	-	-	-	0.972	0.98	0.948	0.971	0.977	0.946
1700	1	IPCW	0.982	0.985	0.977	0.994	0.994	0.987	0.977	0.98	0.965
		GL	-	-	-	0.991	0.993	0.984	0.976	0.978	0.963
	2.5	IPCW	0.975	0.978	0.977	0.984	0.987	0.969	0.96	0.965	0.946
		GL	-	-	-	0.983	0.989	0.969	0.958	0.966	0.946
	5	IPCW	0.985	0.977	0.984	0.983	0.98	0.964	0.968	0.958	0.947
		GL	-	-	-	0.981	0.981	0.96	0.967	0.958	0.946
	10	IPCW	0.978	0.991	0.982	0.982	0.98	0.96	0.96	0.97	0.941
		GL	-	-	-	0.982	0.981	0.961	0.96	0.97	0.941

Table 3.4: Efficiency measure: ratio between sample variance of the main estimators and sample variance of standard estimators. Joint Working models.

Model 1			Censoring aug. estimators			G-estimators			Doubly aug. estimators		
$Time$	k_1	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	1	IPCW	0.995	0.995	0.994	0.989	0.982	0.972	0.983	0.977	0.964
		GL	-	-	-	0.933	0.949	0.884	0.983	0.977	0.964
	2.5	IPCW	0.997	1	0.999	0.971	0.973	0.945	0.969	0.963	0.936
		GL	-	-	-	0.934	0.944	0.88	0.969	0.963	0.936
	5	IPCW	0.997	0.999	0.998	0.959	0.95	0.909	0.95	0.938	0.896
		GL	-	-	-	0.931	0.929	0.861	0.95	0.938	0.896
	10	IPCW	0.999	0.999	0.998	0.955	0.95	0.905	0.942	0.934	0.882
		GL	-	-	-	0.932	0.936	0.868	0.942	0.934	0.882
1700	1	IPCW	0.982	0.985	0.977	0.994	0.997	0.992	0.96	0.962	0.961
		GL	-	-	-	0.799	0.847	0.656	0.959	0.961	0.96
	2.5	IPCW	0.975	0.978	0.977	0.985	0.989	0.974	0.95	0.946	0.925
		GL	-	-	-	0.826	0.851	0.676	0.95	0.946	0.924
	5	IPCW	0.985	0.977	0.984	0.97	0.979	0.95	0.919	0.929	0.888
		GL	-	-	-	0.815	0.868	0.694	0.918	0.929	0.887
	10	IPCW	0.978	0.991	0.982	0.963	0.976	0.939	0.906	0.913	0.862
		GL	-	-	-	0.845	0.885	0.738	0.906	0.913	0.862

Table 3.5: Estimated coverage probability for the estimated standard errors. Stratified working models for Treatment variable.

Model 1			Standard estimators			Censoring aug. estimators			Doubly aug. estimators		
$Time$	k_1	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	1	IPCW	0.949	0.95	0.948	0.948	0.969	0.947	0.95	0.949	0.948
		GL	-	-	-	-	-	-	0.95	0.949	0.948
	2.5	IPCW	0.949	0.951	0.951	0.948	0.97	0.952	0.949	0.953	0.95
		GL	-	-	-	-	-	-	0.949	0.953	0.951
	5	IPCW	0.948	0.949	0.953	0.947	0.968	0.952	0.949	0.951	0.95
		GL	-	-	-	-	-	-	0.949	0.951	0.95
	10	IPCW	0.951	0.953	0.954	0.951	0.971	0.953	0.952	0.954	0.952
		GL	-	-	-	-	-	-	0.952	0.955	0.953
1700	1	IPCW	0.941	0.948	0.949	0.942	0.958	0.949	0.94	0.944	0.948
		GL	-	-	-	-	-	-	0.94	0.944	0.948
	2.5	IPCW	0.948	0.945	0.952	0.943	0.959	0.951	0.944	0.943	0.951
		GL	-	-	-	-	-	-	0.944	0.944	0.95
	5	IPCW	0.943	0.947	0.944	0.939	0.959	0.946	0.939	0.946	0.948
		GL	-	-	-	-	-	-	0.94	0.945	0.948
	10	IPCW	0.946	0.951	0.954	0.94	0.96	0.953	0.939	0.946	0.952
		GL	-	-	-	-	-	-	0.94	0.947	0.952

Table 3.6: Marginal Mean estimations for treated, untreated arms and treatment effect. Stratified working models for Treatment variable.

Model 1			Standard estimators			G-estimators			Doubly aug. estimators		
$Time$	k_1	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	1	IPCW	0.1607	0.1953	-0.0346	0.1607	0.1953	-0.0346	0.1606	0.1951	-0.0345
		GL	-	-	-	0.1606	0.1954	-0.0347	0.1606	0.1951	-0.0345
	2.5	IPCW	0.4019	0.488	-0.0861	0.4019	0.488	-0.0861	0.4015	0.4875	-0.086
		GL	-	-	-	0.4018	0.4883	-0.0865	0.4015	0.4875	-0.086
	5	IPCW	0.8031	0.9761	-0.173	0.8031	0.9762	-0.1731	0.8022	0.9752	-0.173
		GL	-	-	-	0.8029	0.9768	-0.1739	0.8022	0.9752	-0.173
	10	IPCW	1.6059	1.9501	-0.3442	1.606	1.95	-0.344	1.6043	1.948	-0.3437
		GL	-	-	-	1.6062	1.9518	-0.3455	1.6043	1.948	-0.3437
1700	1	IPCW	0.1744	0.2117	-0.0373	0.1744	0.2118	-0.0374	0.1735	0.2106	-0.0371
		GL	-	-	-	0.1742	0.212	-0.0378	0.1735	0.2107	-0.0371
	2.5	IPCW	0.4364	0.529	-0.0927	0.4364	0.529	-0.0926	0.4339	0.5262	-0.0923
		GL	-	-	-	0.4357	0.5298	-0.094	0.4339	0.5262	-0.0923
	5	IPCW	0.872	1.0585	-0.1865	0.872	1.0584	-0.1864	0.8672	1.0527	-0.1855
		GL	-	-	-	0.871	1.0599	-0.1889	0.8672	1.0526	-0.1855
	10	IPCW	1.7428	2.1148	-0.372	1.7428	2.1148	-0.372	1.7332	2.1037	-0.3705
		GL	-	-	-	1.7414	2.1183	-0.3769	1.7332	2.1037	-0.3705

Table 3.7: Efficiency measure: ratio between sample variance of the main estimators and sample variance of standard estimators with simple censoring. Stratified working models for Treatment variable.

Model 1			Censoring aug. estimators			G-estimators			Doubly aug. estimators		
$Time$	k_1	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	1	IPCW	0.976	0.969	0.974	0.994	0.993	0.987	0.97	0.962	0.961
		GL	-	-	-	0.984	0.98	0.977	0.97	0.962	0.961
	2.5	IPCW	0.939	0.943	0.949	0.987	0.988	0.974	0.926	0.93	0.922
		GL	-	-	-	0.956	0.964	0.95	0.925	0.929	0.918
	5	IPCW	0.929	0.906	0.905	0.973	0.976	0.948	0.905	0.884	0.857
		GL	-	-	-	0.937	0.935	0.902	0.903	0.882	0.855
	10	IPCW	0.864	0.857	0.854	0.973	0.973	0.945	0.838	0.832	0.802
		GL	-	-	-	0.919	0.919	0.891	0.834	0.831	0.798
1700	1	IPCW	0.968	0.965	0.962	0.997	1	0.997	0.965	0.966	0.96
		GL	-	-	-	0.99	0.985	0.985	0.967	0.966	0.96
	2.5	IPCW	0.949	0.941	0.942	0.987	0.981	0.968	0.936	0.923	0.911
		GL	-	-	-	0.962	0.954	0.938	0.936	0.923	0.91
	5	IPCW	0.914	0.898	0.911	0.985	0.972	0.955	0.898	0.871	0.865
		GL	-	-	-	0.947	0.933	0.919	0.89	0.87	0.858
	10	IPCW	0.868	0.851	0.854	0.973	0.977	0.949	0.838	0.829	0.802
		GL	-	-	-	0.921	0.922	0.893	0.835	0.823	0.794

Table 3.8: Efficiency measure: ratio between sample variance of the main estimators and sample variance of standard estimators with simple censoring. Joint working models.

Model 1			Censoring aug. estimators			G-estimators			Doubly aug. estimators		
$Time$	k_1	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	1	IPCW	0.976	0.969	0.974	1.018	1.033	1.051	0.965	0.97	0.964
		GL	-	-	-	0.853	0.903	0.773	0.964	0.97	0.963
	2.5	IPCW	0.939	0.943	0.949	1.021	1.037	1.058	0.935	0.934	0.921
		GL	-	-	-	0.868	0.876	0.771	0.935	0.934	0.92
	5	IPCW	0.929	0.906	0.905	1.044	1.05	1.094	0.905	0.893	0.872
		GL	-	-	-	0.871	0.878	0.781	0.904	0.893	0.872
	10	IPCW	0.864	0.857	0.854	1.072	1.085	1.157	0.85	0.842	0.832
		GL	-	-	-	0.851	0.871	0.773	0.849	0.843	0.832
1700	1	IPCW	0.968	0.965	0.962	1.022	1.024	1.045	0.963	0.954	0.947
		GL	-	-	-	0.866	0.894	0.776	0.963	0.954	0.947
	2.5	IPCW	0.949	0.941	0.942	1.041	1.049	1.093	0.926	0.931	0.915
		GL	-	-	-	0.867	0.883	0.763	0.926	0.932	0.915
	5	IPCW	0.914	0.898	0.911	1.063	1.061	1.124	0.895	0.884	0.863
		GL	-	-	-	0.836	0.861	0.739	0.895	0.883	0.862
	10	IPCW	0.868	0.851	0.854	1.084	1.078	1.162	0.841	0.826	0.814
		GL	-	-	-	0.831	0.862	0.748	0.841	0.826	0.815

3.2 Composite event simulations

In this section we will treat the second approach to simulate the data. As mentioned above, it's based on considering cardiovascular deaths as a separate event. Therefore, it has its own process that is a terminal event one, like the non cardiovascular deaths. However, since we consider it as a part of the process of interest, we separated it from the non cardiovascular death, that plays a role of competing event. Therefore, to simulate the data in this setting, we need to estimate three types of cumulative hazard among survivors: $\lambda_r(t)$ as the baseline for the recurrent events (non fatal), $\lambda_{d_1}(t)$ as the baseline for the cardiovascular deaths and $\lambda_{d_2}(t)$ as the baseline referred to the non-cardiovascular ones. Looking at this formulation, the events generated from the first two baselines belong to the process of interest and thus both they contribute to the Marginal means' estimation. Nevertheless, since the cardiovascular deaths are also terminal events, we treated it also as a terminal event. Like in the standard simulations, we will show the $Model1$ setting considering a common frailty variable Z, gamma distributed with variance equal to 1. Therefore, the baselines become $Z\lambda_r(t)$, $Z\lambda_{d_1}(t)$, $Z\lambda_{d_2}(t)$. Moreover, we use the same multiplicative factor k_1 to increase the number of recurrent events, applied only to the first baseline $\lambda_r(t)$. For this reason, we except to see smaller values in marginal mean estimations when $k_1 > 1$ increases respect to the standard simulations, as we can see in Table 3.11. However, we can see that when $k_1 = 1$, marginal means and treatment effect are very similar in the two approaches, see first lines of Tables 3.6 and 3.11. Therefore, both of them are valid in term of bias. About efficiency, looking at Tables 3.9 and 3.10 we can make more or less the same remarks respect the ones made in the previous section with the difference that in general the efficiency gain estimated here is slightly bigger than the one obtained with the standard simulations. About coverage for the estimated standards errors, all the measures are comparable and congruent with the ones reported in Table 3.5.

Table 3.9: Efficiency measure: ratio between sample variance of the main estimators and sample variance of standard estimators. Stratified working model for Treatment variable.

Model 1			Censoring aug. estimators			G-estimators			Doubly aug. estimators		
$Time$	k_1	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	1	IPCW	0.999	0.997	0.998	0.985	0.976	0.96	0.984	0.973	0.958
		GL	-	-	-	0.98	0.971	0.951	0.979	0.97	0.951
	2.5	IPCW	0.996	0.999	1	0.975	0.963	0.939	0.972	0.962	0.939
		GL	-	-	-	0.972	0.962	0.937	0.969	0.962	0.936
	5	IPCW	1	0.997	0.997	0.973	0.962	0.934	0.973	0.958	0.931
		GL	-	-	-	0.972	0.96	0.931	0.973	0.958	0.93
	10	IPCW	0.997	0.998	0.997	0.971	0.973	0.94	0.968	0.97	0.937
		GL	-	-	-	0.968	0.974	0.939	0.967	0.97	0.935
1700	1	IPCW	0.981	0.977	0.98	0.989	0.987	0.976	0.97	0.964	0.955
		GL	-	-	-	0.985	0.983	0.97	0.968	0.965	0.954
	2.5	IPCW	0.972	0.965	0.97	0.992	0.984	0.975	0.963	0.95	0.946
		GL	-	-	-	0.99	0.981	0.971	0.964	0.95	0.946
	5	IPCW	0.964	0.969	0.966	0.984	0.983	0.969	0.949	0.953	0.936
		GL	-	-	-	0.979	0.981	0.965	0.949	0.954	0.937
	10	IPCW	0.958	0.968	0.957	0.979	0.978	0.955	0.938	0.946	0.914
		GL	-	-	-	0.978	0.974	0.951	0.937	0.944	0.912

Table 3.10: Efficiency measure: ratio between sample variance of the main estimators and sample variance of standard estimators. Joint working model.

Model 1			Censoring aug. estimators			G-estimators			Doubly aug. estimators		
$Time$	k_1	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	1	IPCW	0.996	0.996	0.997	0.985	0.979	0.964	0.98	0.974	0.959
		GL	0.996	0.996	0.997	0.933	0.942	0.879	0.98	0.974	0.959
	2.5	IPCW	0.994	0.994	0.994	0.966	0.969	0.937	0.958	0.96	0.925
		GL	0.994	0.994	0.994	0.925	0.929	0.858	0.958	0.96	0.925
	5	IPCW	0.994	0.995	0.994	0.963	0.957	0.92	0.953	0.948	0.906
		GL	0.994	0.995	0.994	0.931	0.934	0.868	0.953	0.948	0.906
	10	IPCW	0.989	0.989	0.99	0.95	0.952	0.903	0.936	0.936	0.885
		GL	0.989	0.989	0.99	0.93	0.93	0.862	0.937	0.936	0.885
1700	1	IPCW	0.971	0.969	0.965	0.993	0.991	0.985	0.959	0.956	0.941
		GL	0.971	0.969	0.965	0.758	0.804	0.572	0.959	0.955	0.939
	2.5	IPCW	0.959	0.955	0.961	0.989	0.989	0.978	0.942	0.937	0.925
		GL	0.959	0.955	0.961	0.779	0.824	0.61	0.941	0.937	0.925
	5	IPCW	0.953	0.94	0.949	0.982	0.98	0.963	0.925	0.909	0.889
		GL	0.953	0.94	0.949	0.796	0.817	0.625	0.924	0.909	0.889
	10	IPCW	0.952	0.945	0.955	0.976	0.987	0.963	0.912	0.91	0.879
		GL	0.952	0.945	0.955	0.826	0.855	0.686	0.912	0.909	0.878

Table 3.11: Marginal Mean estimations for treated, untreated arms and treatment effect.

Model 1			Standard estimators			G-estimators			Doubly aug. estimators		
$Time$	k_1	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	1	IPCW	0.1682	0.2061	-0.0379	0.1682	0.2061	-0.0379	0.168	0.2059	-0.0379
		GL	-	-	-	0.1681	0.2061	-0.038	0.168	0.2059	-0.0379
	2.5	IPCW	0.3494	0.4218	-0.0724	0.3494	0.4218	-0.0724	0.349	0.4213	-0.0723
		GL	-	-	-	0.3493	0.4219	-0.0725	0.349	0.4213	-0.0723
	5	IPCW	0.6513	0.7809	-0.1296	0.6513	0.7809	-0.1297	0.6506	0.7801	-0.1295
		GL	-	-	-	0.6513	0.7812	-0.1299	0.6506	0.7801	-0.1295
	10	IPCW	1.2551	1.4985	-0.2435	1.255	1.4986	-0.2436	1.2537	1.4969	-0.2433
		GL	-	-	-	1.2552	1.4994	-0.2442	1.2537	1.4969	-0.2433
1700	1	IPCW	0.1854	0.2266	-0.0412	0.1854	0.2267	-0.0412	0.1844	0.2253	-0.0409
		GL	-	-	-	0.1852	0.2268	-0.0416	0.1844	0.2253	-0.0409
	2.5	IPCW	0.3847	0.4636	-0.0789	0.3847	0.4636	-0.0789	0.3821	0.4605	-0.0784
		GL	-	-	-	0.3844	0.464	-0.0796	0.3821	0.4605	-0.0784
	5	IPCW	0.7167	0.8572	-0.1405	0.7167	0.8572	-0.1405	0.7118	0.8514	-0.1396
		GL	-	-	-	0.7161	0.858	-0.1419	0.7118	0.8513	-0.1396
	10	IPCW	1.3802	1.6446	-0.2644	1.3802	1.6446	-0.2644	1.3707	1.6335	-0.2628
		GL	-	-	-	1.379	1.6468	-0.2678	1.3707	1.6335	-0.2628

Table 3.12: Efficiency measure: ratio between sample variance of the main estimators and sample variance of standard estimators with simple censoring. Stratified working model for Treatment variable.

Model 1			Standard estimators			G-estimators			Doubly aug. estimators		
$Time$	k_1	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	1	IPCW	0.976	0.968	0.977	0.992	0.992	0.981	0.968	0.96	0.958
		GL	-	-	-	0.977	0.974	0.968	0.966	0.959	0.956
	2.5	IPCW	0.954	0.96	0.957	0.992	0.991	0.981	0.946	0.949	0.937
		GL	-	-	-	0.968	0.969	0.962	0.948	0.946	0.937
	5	IPCW	0.939	0.94	0.943	0.99	0.975	0.964	0.932	0.915	0.91
		GL	-	-	-	0.96	0.947	0.938	0.931	0.915	0.909
	10	IPCW	0.894	0.881	0.89	0.974	0.979	0.951	0.869	0.858	0.84
		GL	-	-	-	0.932	0.928	0.907	0.864	0.855	0.832
1700	1	IPCW	0.975	0.952	0.965	0.991	0.992	0.98	0.966	0.942	0.943
		GL	-	-	-	0.976	0.96	0.955	0.963	0.94	0.94
	2.5	IPCW	0.949	0.937	0.946	0.995	0.99	0.982	0.945	0.926	0.927
		GL	-	-	-	0.97	0.959	0.956	0.944	0.924	0.925
	5	IPCW	0.919	0.903	0.916	0.979	0.985	0.962	0.898	0.888	0.879
		GL	-	-	-	0.94	0.936	0.926	0.898	0.886	0.877
	10	IPCW	0.87	0.847	0.858	0.976	0.972	0.947	0.847	0.819	0.807
		GL	-	-	-	0.932	0.915	0.894	0.843	0.816	0.8

3.3 Baselines' non proportionality simulations

In this section we illustrate that the G-estimator based on the Gosh-Lin model could be severely biased when the model does not hold over the entire time-range. We simulated data as before but now allowed separate baselines of the recurrent events process for treated and non-treated subjects. In Figure 3.2 we show the baselines used in the proportional Cox setting earlier: the green and blue curves. For the simulation in the non-proportionality setting we used instead the red and blue curves for the treated and untreated subjects. We now conducted a simulation as before with 10000 replications and here report only the means of the estimators for the G-estimators, see Table 3.13. All other estimators were unbiased, and the IPCW G-estimator is still unbiased by construction, and also in the finite sample simulation. In contrast, however, we see a large bias for the Ghosh-Lin based G-estimator. This also, as expected, lead to a much smaller efficiency gain for the doubly augmented estimator when based on the Ghosh-Lin working model (not shown).

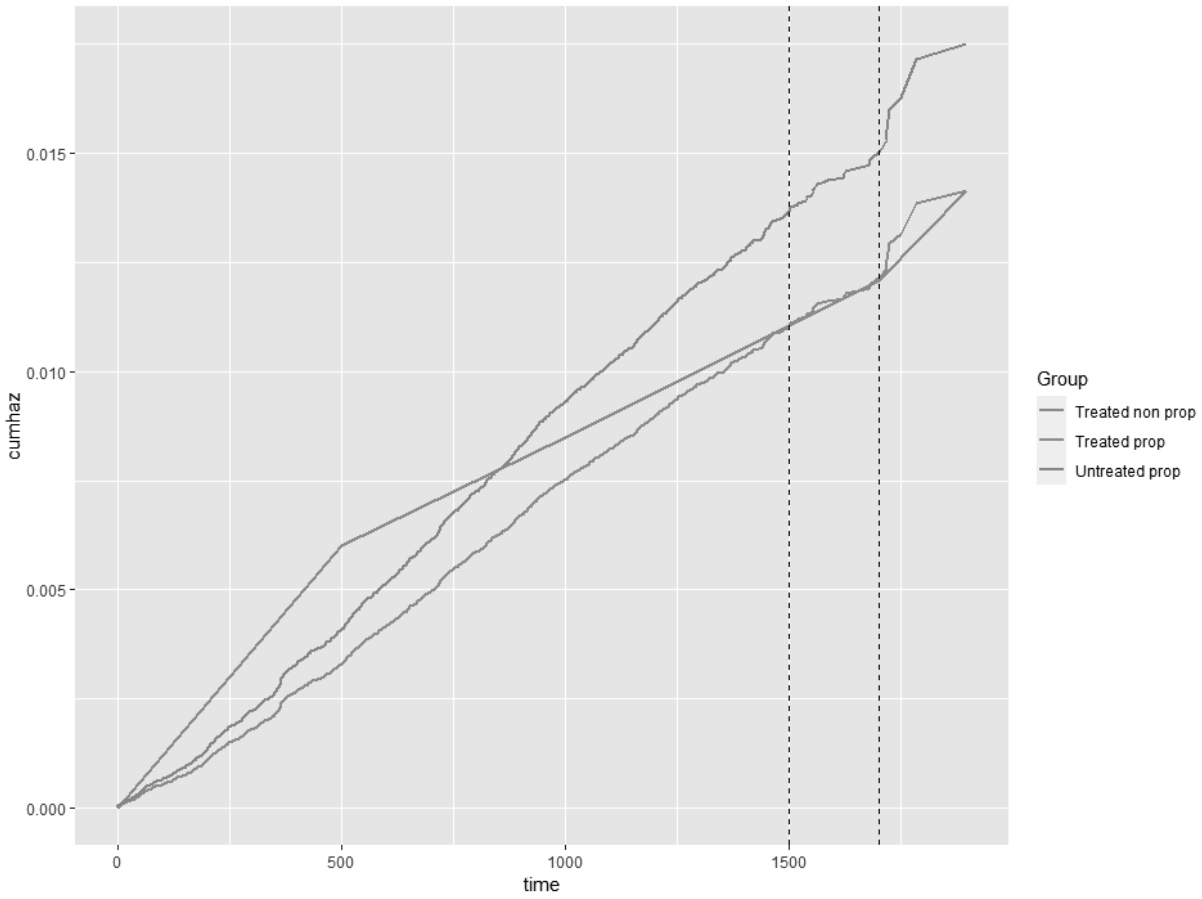

Figure 3.2: Different baselines for Treated and Untreated arms under different hypothesis of proportionality.

Table 3.13: Marginal Mean estimations for treated, untreated arms and treatment effect with non proportional baselines.

		G-estimators		
$Time$	WM	$A=1$	$A=0$	$Diff$
1500	IPCW	0.1696	0.207	-0.0374
	GL	0.2070	0.1924	-0.0080
1700	IPCW	0.1840	0.2255	-0.0415
	GL	0.2005	0.2090	-0.0085

Worked example: LEADER DATA

In this section we present the results related to the original LEADER data. We will show only the results with the stratified working models for the RCT augmentation, since for the original data there are no substantial differences about the efficiency gain, as observed also in the previous simulations when $k_1 = 1$. Moreover, we will show the results about the MI outcome too for completeness, although in the simulations we considered only the 3p-MACE, since it was a more complete and interested type of events. About the estimates of marginal means and treatment effect, they are in accordance with the one obtain in the simulations, with a small difference at the third decimal position between the estimates obtained with the different estimators (standard, censoring and doubly augmented ones). Specially in MI events, we can see that the treatment effect estimate at specific time 1700 is quite bigger for the augmented estimators respect to the standard one (see Table 4.3). Therefore, in this case the augmented estimators tend to overestimate a little the treatment effect. However, we're referring to small differences that probably can not indicate a systematic problem of bias. About the efficiency measure, we can see the best gain is around 96.4%, obtained for the treatment effect estimates with the doubly augmented estimators. Here the measure is the ratio between the estimated variance of the main estimator and the estimated variance of the standard one. We computed such measure for two type of estimators: Censoring augmented and Doubly augmented ones. We can note that the efficiency gain is similar to the one obtained in the simulation with $k_1 = 1$, the setting that simulated the data from the original one. This

also indicates that the gain obtained by simulations is quite realistic.

Table 4.1: Marginal Mean estimations for treated, untreated arms and treatment effect for 3p-MACE events.

		Standard estimator			Censoring aug. estimators			Doubly aug. estimators		
$Time$	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	IPCW	0.1701	0.2032	-0.0331	0.1691	0.2026	-0.0335	0.1675	0.2046	-0.0371
	GL	0.1701	0.2032	-0.0331	0.1691	0.2026	-0.0335	0.1675	0.2047	-0.0372
1700	IPCW	0.1883	0.2143	-0.026	0.1849	0.213	-0.0281	0.1833	0.2149	-0.0316
	GL	0.1883	0.2143	-0.026	0.1849	0.213	-0.0281	0.1831	0.2153	-0.0321

Table 4.2: Efficiency measure: ratio between estimated variance of the main estimators and the estimated variance of the standard ones for 3p-MACE events.

		Censoring aug. estimators			Doubly aug. estimators		
$Time$	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	IPCW	0.994	0.999	0.997	0.982	0.991	0.969
	GL	0.994	0.999	0.997	0.982	0.991	0.969
1700	IPCW	0.982	0.996	0.988	0.975	0.988	0.964
	GL	0.982	0.996	0.988	0.975	0.99	0.964

Table 4.3: Marginal Mean estimations for treated, untreated arms and treatment effect for MI events.

		Standard estimators			Censoring aug. estimators			Doubly aug. estimators		
$Time$	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	IPCW	0.081	0.0937	-0.0126	0.0802	0.0932	-0.013	0.0795	0.0941	-0.0147
	GL	0.081	0.0937	-0.0126	0.0802	0.0932	-0.013	0.0794	0.0942	-0.0148
1700	IPCW	0.0918	0.0973	-0.0054	0.09	0.0964	-0.0064	0.0891	0.0974	-0.0083
	GL	0.0918	0.0973	-0.0054	0.09	0.0964	-0.0064	0.0892	0.0975	-0.0083

Table 4.4: Efficiency measure: ratio between estimated variance of the main estimators and the estimated variance of the standard ones for MI events.

		Censoring aug. estimators			Doubly aug. estimators		
$Time$	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	IPCW	0.989	1	0.994	0.98	0.992	0.977
	GL	0.989	1	0.994	0.98	0.992	0.978
1700	IPCW	0.975	0.998	0.984	0.968	0.988	0.969
	GL	0.975	0.998	0.984	0.969	0.989	0.969

Conclusions and final remarks

In this work we proposed a novel doubly augmented estimator for the marginal mean number of recurrent events. In the first section we showed its theoretical derivation starting from two different influence functions, the one related to the censoring augmented estimator and the one related to the RCT augmentation, permitted by the treatment's randomisation completely known in the context of clinical trials. Moreover, we showed theoretically and asymptotically that we gain in efficiency with respect to the standard estimator, especially if we choose the working model involved in the RCT augmentation in a specific way. Then we also exploited the orthogonality between the censoring-based augmentation and the RCT augmentation.

In the second and third sections, we showed via finite sample simulations and with the worked example about LEADER data, how much we can gain in efficiency in practice, assuming different settings about the dependence between recurrent events and terminal event processes, censoring distribution and rate of recurrent events. In general, we can conclude that the novel doubly augmented estimator provides a substantially improved efficiency in reducing the variance of the estimates in all the settings tested. About the dependence assumption between events and terminal event processes, we saw that the efficiency gain is greater when we do not assume such dependence. However, it's more realistic to consider the case with dependence in practice, since the number of recurrent events is often related to the death time, mostly if the recurrent events are harmful to health like the ones considered in our worked example. Despite this fact, with a quite low number of

recurrent events as in LEADER data (simulations with $k_1 = 1$), there is no so much difference in term of efficiency, independently of the presence of heterogeneity in the data. About the censoring distribution, we simulated data using two different settings: a piecewise constant survival function with a very low censoring rate, which declines by the end of the follow-up; an exponential survival function with a constant censoring rate (Figure 3.1). Regarding the doubly augmented estimator, we obtained a higher gain simulating data from the second setting, thus when the censoring starts at the beginning of the follow-up period. The reason is probably due to the considerable improvement of the censoring augmentation that is transferred also to the doubly augmented estimator. Indeed, when the censoring time starts later, as in LEADER data, the gain reached in the doubly augmented estimator is mostly due to the rct augmentation, since the censoring augmentation does not reveal its potential power.

Indeed, using the simple exponential function as the censoring distribution, the simple IPCW G-estimator performs quite worse in terms of efficiency with respect to the standard estimator. This is not the case of the Gosh-Lin G-estimator, which provides the best gain among all the estimators when we fitted it jointly, although this is not the model suggested by the theory. Therefore, in such situation, we obtain different results about variance reduction. Nevertheless, these differences among the estimators in the different settings do not affect the performance of the doubly augmented estimator that always guaranteed a gain's improvement and it's almost always the estimator that can reach the best gain among the others considered. Indeed, the only estimator that performs better in some cases is the Gosh-Lin G-estimator, although we have to point out that it is a regression model based on more restricted assumptions, not robust in case of misspecification or other violations.

The simulation studies confirm that the performance in reducing the variance provided by the doubly augmented estimator is substantially high in all the several settings considered in the simulations. In

summary, we can conclude that the proposed estimator is more efficient respect to the standard one. The efficiency's amount depends mostly on the number of recurrent events. Indeed, as we saw in the simulations and the worked example, the LEADER data may not be the most appropriate setting to see the potential gain that such estimator would be able to reach, since the number of recurrent events is not so relevant. Other factors that influence the performance of the two augmentations terms in the doubly augmented estimator are the censoring distribution and the randomisation, which guarantees the knowledge of treatment assignment's probability that's independent of the auxiliary covariates X. If the efficiency gain reached by the doubly augmented estimator is high enough, first of all we can obtain more accurate estimates respect to the standard estimator. This consents to making more robust and precise conclusions about the effect of a new drug or treatment in several settings, also avoiding restrictive assumptions on which other models are based. Moreover, by adopting an estimator that is more efficient, the study of interest can reach a higher statistical power, and thus it would need a lower sample size with respect to using standard estimators. The consequence is a reduced effort to follow a larger number of subjects, especially when a long follow-up time is necessary. In regard to this point, the contribution of my thesis answers the important request by the European Medicines Agency about efficient statistical estimation methods for the treatment effect on estimands of recurrent event endpoints (see Akacha et al. (2018)).

Appendix A: Variance computations

A.1: Variance reduction using the censoring augmentation

We'll derive the variance of the augmented estimator from its influence function that's equal to:

$$\int_0^t \frac{Z(s,a)}{\hat{G}_c(s,a)} dN(s,a) + \int_0^t \frac{L^{eff}(s,t,a)}{\hat{G}_c(s,a)} d\hat{M}^C(s) - \mu(t,a)$$
$$= \int_0^t \frac{Z(s,a)}{\hat{G}_c(s,a)} dN(s,a) + \int_0^t \frac{E[H(s,t,a)|\mathcal{H}(s,a), D > s)]}{\hat{G}_c(s,a)} d\hat{M}^C(s) - \mu(t,a)$$
$$= (Z(s,a) - \mu(t,a))$$
$$- \int_0^t [H(s,t,a) - E(H(s,t,a)|\mathcal{H}(s,a), D > s] \frac{Z(s,a)}{\hat{G}_c(s,a)} d\hat{M}^C(s).$$

where $\mathcal{H}(s,1)$ is the history among those treated. The variance of this influence function is

$$E(Z(s,a) - \mu(t,a))^2$$
$$+ E\left(\int_0^t [H(s,t,a) - E(H(s,t,a)|\mathcal{H}(s), D > s]^2 \left(\frac{Z(s,a)}{\hat{G}_c(s,a)} \right)^2 d\hat{M}^C(s) \right)$$
$$= E(Z(s,a) - \mu(t,a))^2 + E\left(\int_0^t (\Delta H)^2 \frac{Z(s,a)}{\hat{G}_c^2(s,a)} d\Lambda_c(s) \right)$$

Now we consider the variance of the standard estimator that considers the simple mean $E(H(s,t,a)|D > s)$ instead of the conditional one

$E(H(s,t,a)|\mathcal{H}(s,a), D > s)$ is equal to:

$$E(Z(s,a) - \mu(t,a))^2 + E\left(\int_0^t [H(s,t,a) - E(H(s,t,a)|D > s]^2 \frac{Z(s,a)}{\hat{G}_c(s,a)} d\Lambda_c(s)\right)$$

$$E(Z(s,a) - \mu(t,a))^2 + E\left(\int_0^t (\Delta_b(H))^2 \frac{Z(s,a)}{\hat{G}_c^2(s,a)} d\Lambda_c(s)\right),$$

where $\Delta_b(H) = H(s,t,a) - E(H(s,t,a)|D > s)$. Focusing on the mean of the integrand, we can write:

$$\begin{aligned}
&E[(\Delta_b H)^2 Z(s,a)] = E[(\Delta_b H)^2 | D > s)] S(s,a) G_c(s,a) \\
&= E\{[H(s,t,a) - E(H(s,t,a)|\mathcal{H}(s,a), D > s) \\
&+ E(H(s,t,a)|\mathcal{H}(s,a), D > s) - E(H(s,t,a)|D > s)]^2 | D > s\} S(s,a) G_c(s,a) \\
&= \{E[H(s,t,a) - E(H(s,t,a)|\mathcal{H}(s,a), D > s)]^2 \\
&+ E[E(H(s,t,a)|\mathcal{H}(s,a), D > s) - E(H(s,t,a)|D > s)]^2 | D > s\} S(s,a) G_c(s,a) \\
&= \{E(\Delta H | D > s)^2 + E[E(H(s,t,a)|\mathcal{H}(s,a), D > s) - \\
&E(H(s,t,a)|D > s)]^2\} S(s,a) G_c(s,a).
\end{aligned}$$

Therefore, the reduction of variance due to the augmentation term is equal to:

$$\begin{aligned}
&\int_0^t \{E[E[H(s,t,a)|\mathcal{H}(s,a), D > s] \\
&- E(H(s,t,a|D > s))]^2 | D > s\} S(s,a) G_c(s,a) d\Lambda_c(s)
\end{aligned}$$

A.2: Variance derivation for the rct augmented estimator ρ_1 (Treatment arm)

The influence function of the efficient estimator for the treated arm is equal to:

$$\frac{A}{\delta}(Z-\rho_1)-\frac{A-\delta}{\delta}(E(Z|A=1,X)-\rho_1)$$

Then the variance can be computed as follows:

$$\begin{aligned}&var\left(\frac{A}{\delta}Z-\frac{A-\delta}{\delta}E(Z|A=1,X)-\rho_{a=1}\right)\\&=var\left(\frac{A}{\delta}(Z-\rho_{a=1})-\frac{A-\delta}{\delta}(E(Z|A=1,X)-\rho_{a=1})\right).\end{aligned}$$

Then, if the working model is correct and thus $E(Z|X,A=1)=F(\beta,X,A=1)$ we can write:

$$\begin{aligned}&var\left(\frac{A}{\delta}Z-\frac{A-\delta}{\delta}F(\beta,X,A=1)-\rho_{a=1}\right)\\&=var\left(\frac{A}{\delta}(Z-\rho_{a=1})-\frac{A-\delta}{\delta}(F(\beta,X,A=1)-\rho_{a=1})\right)\\&=E\left(\frac{A}{\delta}(Z-\rho_{a=1})\right)^2+E\left(\frac{A-\delta}{\delta}(F(\beta,X,A=1)-\rho_{a=1})\right)^2\\&-2E\left(\frac{A}{\delta}(Z-\rho_{a=1})\frac{A-\delta}{\delta}(F(\beta,X,A=1)-\rho_{a=1})\right),\end{aligned}$$

and focusing on the last term (double product), using the conditional expected value property:

$$
\begin{aligned}
&-2E\left(\frac{A}{\delta}(Z-\rho_{a=1})\frac{A-\delta}{\delta}(F(\beta,X,A=1)-\rho_{a=1})\right)\\
&=-2E\left(E\left(\frac{A}{\delta}(Z-\rho_{a=1})\frac{A-\delta}{\delta}(F(\beta,X,A=1)-\rho_{a=1})|X,A\right)\right)\\
&=-2E\Bigg(E\Bigg(\frac{A}{\delta}(Z+F(\beta,X,A=1)-F(\beta,X,A=1)-\rho_{a=1})\\
&\frac{A-\delta}{\delta}(F(\beta,X,A=1)-\rho_{a=1})|X,A\Bigg)\Bigg)\\
&=-2E\Bigg(\frac{A(A-\delta)}{\delta^2}(F(\beta,X,A=1)-\rho_{a=1})\\
&(E(Z-F(\beta,X,A=1)|X,A)+F(\beta,X,A=1)-\rho_{a=1})\Bigg)\\
&=-2E\Bigg(\frac{A(A-\delta)}{\delta^2}(F(\beta,X,A=1)-\rho_{a=1})(F(\beta,X,A=1)-\rho_{a=1})\Bigg)\\
&=-2E\Bigg(\frac{A(A-\delta)}{\delta^2}(F(\beta,X,A=1)-\rho_{a=1})^2\Bigg),
\end{aligned}
$$

because $E(Z-F(\beta,X,A=1)|X,A)=0$, thanks to the conditional mean. Writing $E(\frac{A(A-\delta)}{\delta^2})=E(\frac{A(A-\delta)-\delta(A-\delta)}{\delta^2})=E(\frac{(A-\delta)^2}{\delta^2})$, since that

$$
\begin{aligned}
&E\left(\frac{A(A-\delta)-\delta(A-\delta)}{\delta^2}\right)\\
&=\frac{1}{\delta^2}(E(A(A-\delta))-E(\delta(A-\delta)))\\
&=\frac{1}{\delta^2}(E(A(A-\delta))),
\end{aligned}
$$

because $E(\delta(A-\delta))=\delta^2-\delta^2=0$, we obtain that the double product is equal to:

$$
-2E\left(\frac{(A-\delta)}{\delta}(F(\beta,X,A=1)-\rho_{a=1})\right)^2.
$$

Therefore, the full variance becomes:

$$
\begin{aligned}
&E\left(\frac{A}{\delta}(Z-\rho_1)\right)^2 + E\left(\frac{A-\delta}{\delta}(E(Z|X,A=1)-\rho_1)\right)^2 \\
&- 2E\left(\frac{A(A-\delta)}{\delta^2}(Z-\rho_1)(E(Z|X,A=1)-\rho_1)\right)^2 \\
&= E\left(\frac{A}{\delta}(Z-\rho_1)\right)^2 - E\left(\frac{A-\delta}{\delta}(E(Z|X,X,A=1)-\rho_1)\right)^2.
\end{aligned}
$$

A.3: Variance derivation for the rct augmented estimator ρ_0 (Untreatment arm)

We'll derive the variance of the rct augmented estimator from its influence function that's equal to:

$$
\frac{1-A}{1-\delta}Z - \frac{(1-A)-(1-\delta)}{1-\delta}E(Z|X,A=0) - \rho_{a=0}
$$

, see Tsiatis 2006. Similarly to the statement regarding the Treated arm, taking $F(\beta, X, A=0)$ as working model replacing $E(Z|X,A=0)$, we obtain:

$$
\begin{aligned}
&\frac{1-A}{1-\delta}Z - \frac{(1-A)-(1-\delta)}{1-\delta}F(\hat{\beta},X,A=0) - \rho_{a=0} \\
&= \frac{1-A}{1-\delta}(Z-F(\hat{\beta},X,A=0)) + F(\hat{\beta},X,A=0) - \rho_{a=1},
\end{aligned}
$$

and again if the working model is corrected specified,
$\frac{1-A}{1-\delta}(Z-F(\hat{\beta},X,A=0))=0$ and it is the efficient estimation for $\rho_{a=0}$.

Therefore, the variance of such estimator is equal to:

$$
\begin{aligned}
&var\left(\frac{1-A}{1-\delta}Z - \frac{(1-A)-(1-\delta)}{1-\delta}F(\beta, X, A=0) - \rho_{a=0}\right) \\
&= var\left(\frac{1-A}{1-\delta}(Z-\rho_{a=0}) - \frac{(1-A)-(1-\delta)}{1-\delta}(F(\beta, X, A=0) - \rho_{a=0})\right) \\
&= E\left(\frac{1-A}{1-\delta}(Z-\rho_{a=0})\right)^2 + E\left(\frac{(1-A)-(1-\delta)}{1-\delta}(F(\beta, X, A=0) - \rho_{a=0})\right)^2 \\
&- 2E\left(\frac{1-A}{1-\delta}(Z-\rho_{a=0})\frac{(1-A)-(1-\delta)}{(1-\delta)}(F(\beta, X, A=0) - \rho_{a=0})\right),
\end{aligned}
$$

and focusing on the last term (double product), using the conditional expected value property:

$$
\begin{aligned}
&- 2E\left(\frac{1-A}{1-\delta}(Z-\rho_{a=0})\frac{(1-A)-(1-\delta)}{1-\delta}(F(\beta, X, A=0) - \rho_{a=0})\right) \\
&= -2E\left[E\left(\frac{1-A}{1-\delta}(Z-\rho_{a=0})\frac{(1-A)-(1-\delta)}{1-\delta}(F(\beta, X, A=0) - \rho_{a=0})|X, A\right)\right] \\
&= -2E\left[E\left(\frac{1-A}{1-\delta}(Z + F(\beta, X, A=0) - F(\beta, X, A=0) - \rho_{a=0})\right.\right. \\
&\left.\left.\frac{(1-A)-(1-\delta)}{1-\delta}(F(\beta, X, A=0) - \rho_{a=0})|X, A\right)\right] \\
&= -2E\left(\frac{(1-A)((1-A)-(1-\delta))}{(1-\delta)^2}(F(\beta, X, A=0) - \rho_{a=0})\right. \\
&\left.(E(Z - F(\beta, X, A=0)|X, A) + F(\beta, X, A=0) - \rho_{a=0})\right) \\
&= -2E\left(\frac{(1-A)((1-A)-(1-\delta))}{(1-\delta)^2}(F(\beta, X, A=0) - \rho_{a=0})\right. \\
&\left.(F(\beta, X, A=0) - \rho_{a=0})\right) \\
&= -2E\left(\frac{(1-A)((1-A)-(1-\delta))}{(1-\delta)^2}(F(\beta, X, A=0) - \rho_{a=0})^2\right),
\end{aligned}
$$

since due to the conditional mean $E(Z - F(\beta, X, A = 0)|X, A) = 0$. Also we can write:

$$\begin{aligned}
&E((1-A)((1-A)-(1-\delta))) \\
&= E((1-A)((1-A)-(1-\delta)) - (1-\delta)((1-A)-(1-\delta))) \\
&= E(((1-A)-(1-\delta))^2),
\end{aligned}$$

and

$$\begin{aligned}
&E((1-A)((1-A)-(1-\delta)) - (1-\delta)((1-A)-(1-\delta))) \\
&= E(((1-A)-(1-\delta))^2),
\end{aligned}$$

seeing as $E((1-\delta)((1-A)(1-\delta))) = 0$.

Therefore, the full variance for the untreated arm becomes:

$$\begin{aligned}
&E\left(\frac{1-A}{1-\delta}(Z-\rho_{a=0})\right)^2 + E\left(\frac{(1-A)-(1-\delta)}{1-\delta}(F(\beta,X,A=0)-\rho_{a=0})\right)^2 \\
&- 2E\left(\frac{(1-A)-(1-\delta)}{1-\delta}(F(\beta,X,A=0)-\rho_{a=0})\right)^2 \\
&= E\left(\frac{1-A}{1-\delta}(Z-\rho_{a=0})\right)^2 - E\left(\frac{(1-A)-(1-\delta)}{1-\delta}(F(\beta,X,A=0)-\rho_{a=0})\right)^2
\end{aligned}$$

A.4: Variance for the doubly augmented estimator.

If, rather than Z we consider the IPCW version $\int_0^t \hat{Z}(s,a)\frac{1}{G_c(s,a)}dN(s,a)$ in case of censored subjects, the influence function for the efficient estimator becomes:

$$\frac{A}{\delta}\int_0^t Z(s,a)\frac{1}{G_c(s,a)}dN(s,a) - \frac{A-\delta}{\delta}E(Z|A=1,X) - \rho_{a=1},$$

and if we augment the first term in the way described in section 1.4, we obtain:

$$\begin{aligned}
&\frac{A}{\delta}\Big(\int_0^t Z(s,a)\frac{1}{G_c(s,a)}dN(s,a)\\
&+\int_0^t \frac{E[H_i(s,t,a)|H_i(s,a),D_i>s]}{G_c(s,a)}dM^C(s) - \rho_{a=1}\Big)\\
&-\frac{A-\delta}{\delta}\left(E(Z|A=1,X)-\rho_{a=1}\right).
\end{aligned}$$

Now we can show that the variance of the influence function of such estimator is comparable with the one already derived for the simple observed Z in the previous section. Focusing on the first integral of the augmented estimator, we can write it as:

$$\begin{aligned}
&\int_0^t Z(s,a)\frac{1}{G_c(s,a)}dN(s,a)\\
&=\int_0^t r(s)I(D>s)dN(s) = \int_0^t \frac{I(C>s)}{G_c(s)}I(D>s)dN(s)\\
&=\int_0^t \left(1-\int_0^s \frac{1}{G_c(u)}dM_c(u)\right)I(D>s)dN(s)\\
&=\int_0^t I(D>s)dN(s) - \int_0^t\left(\int_s^t I(D>u)dN(u)\right)\frac{1}{G_c(s)}dM_c(s)\\
&=\int_0^t I(D>s)dN(s) - \int_0^t H(s,t)\frac{1}{G_c(s)}dM_c(s)\\
&=Z-\int_0^t H(s,t)\frac{1}{G_c(s)}dM_c(s),
\end{aligned}$$

changing the order of integration and using $H(s,t) = \int_s^t I(D > u)dN(u)$ and the fact that $\int_0^t I(D > s)dN(s)$ is equal to Z, the original number of recurrent event observed without censored subjects (see Cortese and Scheike 2022, Blanche et al. 2022). From this last formulation, we can rewrite the influence function for the double augmented estimator, the most efficient one, we obtain:

$$\frac{A}{\delta}(Z - \int_0^t (H(s,t) - E[H_i(s,t,a)|H_i(s,a), D_i > s])\frac{1}{G_c(s)}dM_c(s) - \rho_{a=1}) - \frac{A-\delta}{\delta}(E(Z|A=1,X) - \rho_{a=1}).$$

From this result, we can note that the integral is referred to the Martingale measure and thus it has mean zero. Therefore, the variance of the double augmented estimator is equal to the one proposed for the original observed value Z derived in Appendix A.2. Obviously, the same arguments can be done for the double augmented estimators for the Untreated arm. In conclusion, if we substitute the theoretical quantities with the estimated ones the double augmented estimator becomes:

$$\hat{\rho}_{a=1} = \tilde{\mu}(t, a=1) - \frac{1}{n}\sum_i^n \frac{A_i - \delta}{\delta}E(Z_i|X_i, A=1),$$

$$\hat{\rho}_{a=0} = \tilde{\mu}(t, a=0) + \frac{1}{n}\sum_i^n \frac{A_i - \delta}{\delta}E(Z_i|X_i, A=0),$$

where the working model can be either the standard IPCW or the Gosh-Lin one.

A.5: Variance for treatment difference.

We look for best function, $h^*(X)$ for

$$\begin{aligned}\psi &= \frac{A}{\delta}(Z-\rho_1) - \frac{1-A}{1-\delta}(Z-\rho_0) - (A-\delta)(h^*(X)-\tilde{\nu}) \\ &= (Z(A)-\nu(A)) - (A-\delta)(h^*(X)-\tilde{\nu})\end{aligned}$$

with $Z(A) = AZ/\delta - (1-A)Z/(1-\delta) = Z(A-\delta)/(\delta(1-\delta))$, $\nu(A) = A\rho_1/\delta - (1-A)\rho_0/(1-\delta)$, and $\tilde{\nu} = \rho_1/\delta + \rho_0/(1-\delta)$. Let further $E_j(X) = E(Z|X, A=j)$ and note that

$$\begin{aligned}E(Z(A)(A-\delta)|X) &= \delta(1-\delta)\left(\frac{E_1(X)}{\delta} + \frac{E_0(X)}{1-\delta}\right) \\ E(\nu(A)(A-\delta)|X) &= \delta(1-\delta)\left(\frac{\delta_1}{\delta} + \frac{\delta_0}{1-\delta}\right).\end{aligned}$$

Now simply calculating we get

$$\begin{aligned}&E(\psi^2) \\ &= c_1 + E((A-\delta)^2(h^*(X)-\tilde{\nu})^2) - 2E([Z(A)-\nu(A)]\,(A-\delta)(h^*(X)-\tilde{\nu})) \\ &= c_1 + \delta(1-\delta)\Bigg(E((h^*(X)-\tilde{\nu})^2) - 2E\Bigg(\left[\frac{1}{\delta}(E_1(X)-\rho_1) + \frac{1}{1-\delta}(E_0(X)-\rho_0)\right] \\ &(h^*(X)-\tilde{\nu})\Bigg)\Bigg) \\ &= c_1 + \delta(1-\delta)\left(E((h^*(X)-\tilde{\nu})^2) - 2E\left(\left[\frac{E_1(X)}{\delta} + \frac{E_0(X)}{1-\delta} - \nu\right](h^*(X)-\tilde{\nu})\right)\right) \\ &= c_2 + \delta(1-\delta)E\left(h^*(X) - \left[\frac{E_1(X)}{\delta} + \frac{E_0(X)}{1-\delta}\right]\right)^2 \\ &= c_3 + \delta(1-\delta)E\left(h^*(X) - \tilde{Z}(A)\right)^2\end{aligned}$$

with $\tilde{Z}(A) = (AZ)/\delta^2 + ((1-A)Z)/(1-\delta)^2$ and therfore the optimal $h^*(X) = E_1(X)/\delta + E_0(X)/(1-\delta)$. The mean of $\tilde{Z}(A)$ is $E(\tilde{Z}(A)|X) = E_1(X)/\delta + E_0(X)/(1-\delta)$. Note that when using the optimal $h^*(X)$ we note that this is indeed the difference of the two efficient influence functions for the two means of each group.

A.6: Variance of the efficient influence function for the rct augmented estimator using working model (Treatment arm)

Calculating the variance of the influence function directly for a given working model we get

$$
\begin{aligned}
&E\left[\left(\frac{A}{\delta}(Z-\rho_1)-\frac{(A-\delta)}{\delta}(h^*(X)-\rho_1)\right)^2\right]\\
&=E\Bigg[\Bigg(\frac{A}{\delta}(Z-E(Z|X;A=1))+\frac{A}{\delta}(E(Z|X;A=1)-\rho_1)\\
&-\frac{(A-\delta)}{\delta}(h^*(X)-\rho_1)\Bigg)^2\Bigg]\\
&=E\left(\frac{A}{\delta}(Z-E(Z|X;A=1))\right)^2+\\
&E\left(\frac{A}{\delta}(E(Z|X;A=1)-\rho_1)-\frac{(A-\delta)}{\delta}(h^*(X)-\rho_1)\right)^2\\
&=c_1+E\left(\frac{(A-\delta)^2}{\delta^2}(h^*(X)-\rho_1)^2\right)\\
&-2E\left(\frac{A}{\delta}\frac{(A-\delta)}{\delta}(E(Z|X;A=1)-\rho_1)(h^*(X)-\rho_1)\right)\\
&=c_1+\frac{1-\delta}{\delta}\left[E\left((h^*(X)-\rho_1)^2\right)-2E\left(E(Z|X;A=1)-\rho_1)(h^*(X)-\rho_1)\right)\right]\\
&=c_2+\frac{1-\delta}{\delta}E\left(h^*(X)-E(Z|X;A=1)\right)^2\\
&=c_2+\frac{1-\delta}{\delta}E\left(\frac{A}{\delta}(h^*(X)-E(Z|X;A=1))^2\right)\\
&=c_3+\frac{1-\delta}{\delta}E\left(\frac{A}{\delta}(Z-h^*(X))^2\right)\\
&=c_3+\frac{1-\delta}{\delta}E\left(\frac{A}{\delta}\left[(Z-E(Z|X,A=1))+(E(Z|X,A=1)-h^*(X))\right]^2\right)
\end{aligned}
$$

using again the independence between A and Z and where constants do not depend on h^*.

We note also that choosing $h^*(X) = \rho_1$ then we get the simple unadjusted estimator based on the mean of the treatment group.

Appendix B: Additional Simulations for Model 0

B.1: Standard simulations with $Model0$ case.

Table B.1: Efficiency measure: ratio between sample variance of the main estimators and sample variance of standard estimators. Stratified working models for Treatment variable.

Model 0			Censoring aug. estimators			G-estimators			Doubly aug. estimators		
$Time$	k_1	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	1	IPCW	1	0.998	0.999	0.984	0.977	0.959	0.984	0.975	0.958
		GL	-	-	-	0.985	0.974	0.956	0.984	0.973	0.956
	2.5	IPCW	0.995	0.998	0.996	0.967	0.959	0.924	0.963	0.957	0.92
		GL	-	-	-	0.963	0.955	0.916	0.96	0.955	0.914
	5	IPCW	0.999	0.996	0.998	0.948	0.954	0.901	0.947	0.95	0.899
		GL	-	-	-	0.95	0.954	0.904	0.948	0.951	0.902
	10	IPCW	0.994	0.993	0.994	0.944	0.935	0.877	0.938	0.928	0.871
		GL	-	-	-	0.944	0.934	0.876	0.939	0.927	0.87
1700	1	IPCW	0.985	0.976	0.984	0.988	0.986	0.974	0.973	0.962	0.958
		GL	-	-	-	0.985	0.986	0.972	0.973	0.961	0.956
	2.5	IPCW	0.976	0.98	0.974	0.982	0.972	0.952	0.958	0.953	0.928
		GL	-	-	-	0.975	0.969	0.946	0.955	0.952	0.925
	5	IPCW	0.972	0.968	0.973	0.957	0.966	0.924	0.93	0.934	0.897
		GL	-	-	-	0.955	0.965	0.923	0.93	0.934	0.897
	10	IPCW	0.978	0.978	0.983	0.961	0.967	0.928	0.94	0.947	0.914
		GL	-	-	-	0.964	0.971	0.932	0.943	0.947	0.917

Table B.2: Efficiency measure: ratio between sample variance of the main estimators and sample variance of standard estimators.

Model 0			Censoring aug. estimators			G-estimators			Doubly aug. estimators		
$Time$	k_1	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	1	IPCW	0.997	0.997	0.996	0.979	0.978	0.956	0.976	0.973	0.951
		GL	0.997	0.997	0.996	0.928	0.934	0.863	0.975	0.973	0.951
	2.5	IPCW	0.997	0.996	0.996	0.96	0.949	0.909	0.951	0.942	0.898
		GL	0.997	0.996	0.996	0.912	0.91	0.826	0.951	0.942	0.898
	5	IPCW	0.995	0.996	0.995	0.92	0.914	0.837	0.91	0.904	0.821
		GL	0.995	0.996	0.995	0.872	0.869	0.748	0.91	0.904	0.821
	10	IPCW	0.993	0.991	0.992	0.875	0.861	0.738	0.858	0.845	0.713
		GL	0.993	0.991	0.992	0.823	0.816	0.648	0.858	0.845	0.713
1700	1	IPCW	0.979	0.98	0.975	0.991	0.994	0.985	0.968	0.968	0.953
		GL	0.979	0.98	0.975	0.774	0.825	0.604	0.968	0.968	0.952
	2.5	IPCW	0.973	0.978	0.978	0.975	0.979	0.954	0.937	0.946	0.909
		GL	0.973	0.978	0.978	0.75	0.805	0.567	0.936	0.946	0.908
	5	IPCW	0.967	0.969	0.967	0.964	0.96	0.924	0.913	0.91	0.853
		GL	0.967	0.969	0.967	0.743	0.781	0.54	0.914	0.909	0.852
	10	IPCW	0.961	0.959	0.959	0.932	0.924	0.858	0.869	0.846	0.757
		GL	0.961	0.959	0.959	0.728	0.733	0.486	0.869	0.846	0.756

Table B.3: Estimated coverage probability for the estimated standard errors. Stratified working models for Treatment variable.

Model 0			Standard estimators			Censoring aug. estimators			Doubly aug. estimators		
$Time$	k_1	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	1	IPCW	0.947	0.948	0.949	0.946	0.968	0.949	0.949	0.95	0.949
		GL	-	-	-	-	-	-	0.949	0.95	0.95
	2.5	IPCW	0.947	0.948	0.948	0.947	0.969	0.949	0.945	0.949	0.949
		GL	-	-	-	-	-	-	0.945	0.949	0.948
	5	IPCW	0.944	0.947	0.946	0.944	0.97	0.945	0.945	0.95	0.943
		GL	-	-	-	-	-	-	0.944	0.95	0.944
	10	IPCW	0.947	0.947	0.946	0.946	0.97	0.947	0.95	0.954	0.945
		GL	-	-	-	-	-	-	0.95	0.954	0.945
1700	1	IPCW	0.942	0.947	0.95	0.94	0.959	0.95	0.939	0.945	0.948
		GL	-	-	-	-	-	-	0.939	0.945	0.949
	2.5	IPCW	0.946	0.949	0.949	0.941	0.962	0.948	0.943	0.944	0.949
		GL	-	-	-	-	-	-	0.943	0.944	0.95
	5	IPCW	0.951	0.947	0.949	0.948	0.956	0.949	0.945	0.944	0.948
		GL	-	-	-	-	-	-	0.946	0.945	0.947
	10	IPCW	0.95	0.95	0.948	0.946	0.959	0.95	0.944	0.95	0.946
		GL	-	-	-	-	-	-	0.943	0.95	0.946

Table B.4: Marginal Mean estimations for treated, untreated arms and treatment effect. Stratified working models for Treatment variable.

Model 0			Standard estimators			G-estimators			Doubly aug. estimators		
$Time$	k_1	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	1	IPCW	0.1679	0.2057	-0.0378	0.1679	0.2057	-0.0378	0.1678	0.2056	-0.0378
		GL	-	-	-	0.1678	0.2058	-0.0379	0.1678	0.2056	-0.0378
	2.5	IPCW	0.4199	0.5141	-0.0942	0.4199	0.5141	-0.0942	0.4196	0.5137	-0.0941
		GL	-	-	-	0.4198	0.5143	-0.0945	0.4196	0.5137	-0.0941
	5	IPCW	0.8393	1.0281	-0.1888	0.8393	1.0282	-0.1889	0.8387	1.0273	-0.1887
		GL	-	-	-	0.8392	1.0287	-0.1894	0.8387	1.0273	-0.1887
	10	IPCW	1.6788	2.0554	-0.3766	1.6787	2.0555	-0.3769	1.6774	2.054	-0.3766
		GL	-	-	-	1.6791	2.0573	-0.3782	1.6774	2.054	-0.3766
1700	1	IPCW	0.183	0.2241	-0.0411	0.183	0.2241	-0.0411	0.1824	0.2233	-0.0409
		GL	-	-	-	0.1829	0.2243	-0.0414	0.1824	0.2233	-0.0409
	2.5	IPCW	0.458	0.5605	-0.1024	0.4581	0.5604	-0.1024	0.4561	0.5581	-0.1019
		GL	-	-	-	0.4577	0.5609	-0.1032	0.4561	0.5581	-0.1019
	5	IPCW	0.9162	1.1202	-0.204	0.9162	1.1202	-0.204	0.9123	1.1155	-0.2032
		GL	-	-	-	0.9155	1.1215	-0.206	0.9123	1.1155	-0.2032
	10	IPCW	1.8312	2.2395	-0.4082	1.8312	2.2396	-0.4084	1.8236	2.2302	-0.4067
		GL	-	-	-	1.8303	2.2427	-0.4124	1.8236	2.2303	-0.4067

Table B.5: Efficiency measure: ratio between sample variance of the main estimators and sample variance of standard estimators with simple censoring. Stratified working model for Treatment variable.

Model 0			Censoring aug. estimators			G-estimators			Doubly aug. estimators		
$Time$	k_1	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	1	IPCW	0.981	0.981	0.979	0.996	0.995	0.989	0.977	0.976	0.969
		GL	-	-	-	0.984	0.982	0.977	0.977	0.975	0.968
	2.5	IPCW	0.967	0.96	0.962	0.978	0.973	0.95	0.946	0.933	0.911
		GL	-	-	-	0.949	0.94	0.917	0.944	0.931	0.907
	5	IPCW	0.925	0.915	0.925	0.971	0.959	0.928	0.896	0.873	0.85
		GL	-	-	-	0.893	0.872	0.84	0.889	0.869	0.842
	10	IPCW	0.899	0.864	0.873	0.928	0.926	0.856	0.83	0.795	0.737
		GL	-	-	-	0.819	0.784	0.72	0.824	0.788	0.727
1700	1	IPCW	0.984	0.975	0.98	0.991	0.992	0.983	0.975	0.968	0.965
		GL	-	-	-	0.983	0.976	0.97	0.975	0.968	0.964
	2.5	IPCW	0.947	0.948	0.944	0.979	0.978	0.955	0.926	0.926	0.899
		GL	-	-	-	0.932	0.934	0.905	0.925	0.924	0.897
	5	IPCW	0.93	0.896	0.916	0.962	0.961	0.919	0.891	0.859	0.835
		GL	-	-	-	0.89	0.857	0.826	0.888	0.854	0.827
	10	IPCW	0.863	0.854	0.855	0.934	0.948	0.884	0.799	0.803	0.74
		GL	-	-	-	0.791	0.799	0.727	0.79	0.801	0.729

Table B.6: Efficiency measure: ratio between sample variance of the main estimators and sample variance of standard estimators with simple censoring. Joint working model.

Model 0			Censoring aug. estimators			G-estimators			Doubly aug. estimators		
$Time$	k_1	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	1	IPCW	0.979	0.969	0.973	1.026	1.024	1.05	0.971	0.959	0.954
		GL	0.979	0.969	0.973	0.858	0.862	0.747	0.972	0.958	0.954
	2.5	IPCW	0.965	0.949	0.953	1.045	1.044	1.088	0.946	0.918	0.903
		GL	0.965	0.949	0.953	0.836	0.826	0.704	0.945	0.917	0.901
	5	IPCW	0.925	0.919	0.917	1.08	1.112	1.192	0.887	0.889	0.849
		GL	0.925	0.919	0.917	0.772	0.802	0.645	0.887	0.889	0.849
	10	IPCW	0.876	0.866	0.87	1.092	1.128	1.219	0.8	0.787	0.717
		GL	0.876	0.866	0.87	0.697	0.714	0.549	0.799	0.787	0.716
1700	1	IPCW	0.979	0.971	0.975	1.019	1.03	1.049	0.972	0.965	0.961
		GL	0.979	0.971	0.975	0.856	0.878	0.755	0.972	0.964	0.96
	2.5	IPCW	0.948	0.951	0.944	1.058	1.064	1.121	0.927	0.925	0.897
		GL	0.948	0.951	0.944	0.818	0.84	0.699	0.925	0.925	0.896
	5	IPCW	0.928	0.909	0.92	1.055	1.092	1.145	0.878	0.862	0.825
		GL	0.928	0.909	0.92	0.765	0.782	0.64	0.878	0.86	0.823
	10	IPCW	0.878	0.862	0.871	1.159	1.182	1.336	0.812	0.788	0.734
		GL	0.878	0.862	0.871	0.702	0.708	0.552	0.812	0.787	0.733

B.2: Composite simulations with $Model0$ case.

Table B.7: Efficiency measure: ratio between sample variance of the main estimators and sample variance of standard estimators. Stratified working models for Treatment variable.

Model 0			Censoring aug. estimators			G-estimators			Doubly aug. estimators		
$Time$	k_1	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	1	IPCW	0.998	0.997	0.998	0.99	0.977	0.965	0.988	0.975	0.963
		GL	-	-	-	0.988	0.974	0.96	0.987	0.973	0.959
	2.5	IPCW	0.997	0.996	0.995	0.967	0.957	0.923	0.964	0.954	0.918
		GL	-	-	-	0.962	0.951	0.914	0.961	0.948	0.91
	5	IPCW	0.999	0.997	0.999	0.951	0.947	0.896	0.95	0.944	0.894
		GL	-	-	-	0.949	0.946	0.891	0.947	0.943	0.891
	10	IPCW	0.994	0.998	0.997	0.931	0.937	0.867	0.926	0.935	0.865
		GL	-	-	-	0.928	0.94	0.868	0.924	0.939	0.867
1700	1	IPCW	0.979	0.972	0.981	0.992	0.988	0.979	0.971	0.961	0.961
		GL	-	-	-	0.985	0.983	0.972	0.969	0.96	0.958
	2.5	IPCW	0.976	0.976	0.972	0.984	0.97	0.953	0.959	0.945	0.924
		GL	-	-	-	0.977	0.966	0.946	0.958	0.944	0.922
	5	IPCW	0.97	0.969	0.964	0.97	0.962	0.932	0.94	0.931	0.897
		GL	-	-	-	0.97	0.954	0.927	0.939	0.929	0.894
	10	IPCW	0.976	0.977	0.977	0.958	0.961	0.917	0.934	0.936	0.893
		GL	-	-	-	0.96	0.959	0.92	0.934	0.937	0.895

Table B.8: Efficiency measure: ratio between sample variance of the main estimators and sample variance of standard estimators. Joint working models.

Model 0			Censoring aug. estimators			G-estimators			Doubly aug. estimators		
$Time$	k_1	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	1	IPCW	0.997	0.997	0.998	0.983	0.976	0.958	0.979	0.97	0.952
		GL	0.997	0.997	0.998	0.932	0.935	0.869	0.979	0.97	0.952
	2.5	IPCW	0.996	0.996	0.995	0.967	0.957	0.924	0.961	0.949	0.914
		GL	0.996	0.996	0.995	0.905	0.914	0.825	0.961	0.949	0.914
	5	IPCW	0.997	0.995	0.996	0.951	0.939	0.887	0.941	0.928	0.871
		GL	0.997	0.995	0.996	0.896	0.894	0.788	0.941	0.928	0.871
	10	IPCW	0.995	0.994	0.994	0.909	0.89	0.798	0.9	0.875	0.778
		GL	0.995	0.994	0.994	0.859	0.851	0.711	0.9	0.875	0.778
1700	1	IPCW	0.974	0.976	0.975	0.994	0.99	0.985	0.963	0.96	0.949
		GL	0.974	0.976	0.975	0.744	0.79	0.548	0.962	0.96	0.948
	2.5	IPCW	0.971	0.971	0.97	0.985	0.985	0.97	0.945	0.943	0.917
		GL	0.971	0.971	0.97	0.74	0.778	0.528	0.945	0.943	0.917
	5	IPCW	0.968	0.968	0.969	0.975	0.981	0.956	0.927	0.928	0.888
		GL	0.968	0.968	0.969	0.724	0.784	0.524	0.927	0.928	0.888
	10	IPCW	0.957	0.955	0.954	0.956	0.956	0.912	0.888	0.882	0.812
		GL	0.957	0.955	0.954	0.717	0.747	0.486	0.888	0.882	0.812

Table B.9: Marginal Mean estimations for treated, untreated arms and treatment effect. Stratified working model for Treatment variable.

Model 0			Standard estimators			G-estimators			Doubly augmented estimators		
$Time$	k_1	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	1	IPCW	0.1683	0.206	-0.0378	0.1683	0.206	-0.0378	0.1681	0.2059	-0.0378
		GL	-	-	-	0.1682	0.2061	-0.0378	0.1681	0.2059	-0.0378
	2.5	IPCW	0.3496	0.4211	-0.0716	0.3495	0.4212	-0.0717	0.3493	0.4209	-0.0716
		GL	-	-	-	0.3495	0.4213	-0.0718	0.3493	0.4209	-0.0716
	5	IPCW	0.6515	0.7809	-0.1294	0.6515	0.7809	-0.1295	0.651	0.7804	-0.1294
		GL	-	-	-	0.6514	0.7812	-0.1298	0.651	0.7804	-0.1294
	10	IPCW	1.2552	1.499	-0.2437	1.2552	1.4991	-0.2439	1.2543	1.498	-0.2437
		GL	-	-	-	1.2554	1.4999	-0.2444	1.2544	1.498	-0.2437
1700	1	IPCW	0.1852	0.2267	-0.0415	0.1852	0.2267	-0.0415	0.1845	0.2258	-0.0413
		GL	-	-	-	0.1851	0.2268	-0.0416	0.1845	0.2258	-0.0413
	2.5	IPCW	0.385	0.4635	-0.0786	0.385	0.4635	-0.0785	0.3832	0.4614	-0.0782
		GL	-	-	-	0.3845	0.464	-0.0794	0.3832	0.4614	-0.0782
	5	IPCW	0.7169	0.8574	-0.1405	0.7169	0.8573	-0.1404	0.7136	0.8533	-0.1397
		GL	-	-	-	0.7161	0.8584	-0.1422	0.7136	0.8533	-0.1397
	10	IPCW	1.3806	1.6458	-0.2653	1.3806	1.6458	-0.2652	1.3741	1.6381	-0.264
		GL	-	-	-	1.3797	1.6475	-0.2678	1.3741	1.6381	-0.264

Table B.10: Efficiency measure: ratio between sample variance of the main estimators and sample variance of standard estimators with simple censoring. Stratified working model for Treatment variable.

Model 0			Censoring aug. estimators			G-estimators			Doubly aug. estimators		
$Time$	k_1	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	1	IPCW	0.982	0.981	0.978	0.991	0.988	0.979	0.973	0.97	0.959
		GL	-	-	-	0.978	0.977	0.965	0.973	0.971	0.959
	2.5	IPCW	0.963	0.963	0.962	0.986	0.982	0.965	0.95	0.945	0.929
		GL	-	-	-	0.959	0.949	0.937	0.95	0.943	0.928
	5	IPCW	0.947	0.944	0.955	0.97	0.96	0.927	0.916	0.903	0.881
		GL	-	-	-	0.915	0.906	0.877	0.911	0.902	0.874
	10	IPCW	0.922	0.896	0.904	0.946	0.941	0.882	0.869	0.839	0.789
		GL	-	-	-	0.869	0.836	0.782	0.867	0.835	0.784
1700	1	IPCW	0.975	0.969	0.972	0.994	0.991	0.983	0.969	0.961	0.955
		GL	-	-	-	0.976	0.972	0.963	0.967	0.961	0.954
	2.5	IPCW	0.976	0.951	0.964	0.983	0.973	0.955	0.958	0.925	0.92
		GL	-	-	-	0.967	0.924	0.918	0.959	0.923	0.918
	5	IPCW	0.939	0.923	0.943	0.968	0.967	0.931	0.908	0.89	0.874
		GL	-	-	-	0.907	0.894	0.872	0.904	0.888	0.869
	10	IPCW	0.906	0.875	0.887	0.956	0.953	0.905	0.865	0.829	0.796
		GL	-	-	-	0.864	0.832	0.795	0.861	0.827	0.791

Table B.11: Efficiency measure: ratio between sample variance of the main estimators and sample variance of standard estimators with simple censoring. Joint working model.

Model 0			Censoring aug. estimators			G-estimators			Doubly aug. estimators		
$Time$	k_1	WM	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$	$A=1$	$A=0$	$Diff$
1500	1	IPCW	0.982	0.981	0.978	0.991	0.988	0.979	0.973	0.97	0.959
		GL	-	-	-	0.978	0.977	0.965	0.973	0.971	0.959
	2.5	IPCW	0.963	0.963	0.962	0.986	0.982	0.965	0.95	0.945	0.929
		GL	-	-	-	0.959	0.949	0.937	0.95	0.943	0.928
	5	IPCW	0.947	0.944	0.955	0.97	0.96	0.927	0.916	0.903	0.881
		GL	-	-	-	0.915	0.906	0.877	0.911	0.902	0.874
	10	IPCW	0.922	0.896	0.904	0.946	0.941	0.882	0.869	0.839	0.789
		GL	-	-	-	0.869	0.836	0.782	0.867	0.835	0.784
1700	1	IPCW	0.975	0.969	0.972	0.994	0.991	0.983	0.969	0.961	0.955
		GL	-	-	-	0.976	0.972	0.963	0.967	0.961	0.954
	2.5	IPCW	0.976	0.951	0.964	0.983	0.973	0.955	0.958	0.925	0.92
		GL	-	-	-	0.967	0.924	0.918	0.959	0.923	0.918
	5	IPCW	0.939	0.923	0.943	0.968	0.967	0.931	0.908	0.89	0.874
		GL	-	-	-	0.907	0.894	0.872	0.904	0.888	0.869
	10	IPCW	0.906	0.875	0.887	0.956	0.953	0.905	0.865	0.829	0.796
		GL	-	-	-	0.864	0.832	0.795	0.861	0.827	0.791

Bibliography

Akacha, M., Binkowitz, B., Bretz, F., Fritsch, A., Hougaard, P., Jahn-Eimermacher, A., Mendolia, F., Ravn, H., Roger, J., Schlömer, P., et al. (2018). Request for chmp qualification opinion: Clinically interpretable treatment effect measures based on recurrent event endpoints that allow for efficient statistical analyses. *Recurrent Event Qualification Opinion Consortium*.

Bang, H. and Tsiatis, A. A. (2000). Estimating medical costs with censored data. *Biometrika*, 87:329–343.

Blanche, P. F., Holt, A., and Scheike, T. (2022). On logistic regression with right censored data, with or without competing risks, and its use for estimating treatment effects. *Lifetime data analysis*.

Cook, R. J. and Lawless, J. F. (1997). Marginal analysis of recurrent events and a terminating event. *Statistics in medicine*, 16(8):911–924.

Cook, R. J. and Lawless, J. F. (2002). Analysis of repeated events. *Statistical Methods in Medical Research*, 11(2):141–166.

Cook, R. J., Lawless, J. F., Lakhal-Chaieb, L., and Lee, K.-A. (2009). Robust estimation of mean functions and treatment effects for recurrent events under event-dependent censoring and termination: Application to skelet al complications in cancer metastatic to bone. *Journal of the American Statistical Association*, 104(485):60–75.

Cortese, G. and Scheike, T. H. (2022). Efficient estimation of the marginal mean of recurrent events. *Journal of the Royal Statistical Society Series C*, 71(5):1787–1821.

Fritsch, A., Schlömer, P., Mendolia, F., Mütze, T., Jahn-Eimermacher, A., and Consortium, R. E. Q. O. (2021). Efficiency comparison of analysis methods for recurrent event and time-to-first event endpoints in the presence of terminal events—application to clinical trials in chronic heart failure. *Statistics in Biopharmaceutical Research*, pages 1–12.

Furberg, J. K., Rasmussen, S., Andersen, P. K., and Ravn, H. (2022). Methodological challenges in the analysis of recurrent events for randomised controlled trials with application to cardiovascular events in leader. *Pharmaceutical Statistics*, 21(1):241–267.

Ghosh, D. and Lin, D. (2000). Nonparametric analysis of recurrent events and death. *Biometrics*, 56(2):554–562.

Ghosh, D. and Lin, D. (2002). Marginal regression models for recurrent and terminal events. *Statistica Sinica*, pages 663–688.

Mao, L. and Lin, D. Y. (2016). Semiparametric regression for the weighted composite endpoint of recurrent and terminal events. *Biostatistics*, 17(2):390–403.

Marso, S. P., Daniels, G. H., Brown-Frandsen, K., Kristensen, P., Mann, J. F., Nauck, M. A., Nissen, S. E., Pocock, S., Poulter, N. R., Ravn, L. S., et al. (2016). Liraglutide and cardiovascular outcomes in type 2 diabetes. *New England Journal of Medicine*, 375(4):311–322.

Robins, J. M. and Rotnitzky, A. (1992). Recovery of information and adjustment of dependent censoring using surrogate markers. In *AIDS Epidemiology-Methodological Issues*, pages 24–33. Birkhäuser, Boston.

Rosenblum, M. and Steingrimsson, J. A. (2016). Matching the efficiency gains of the logistic regression estimator while avoiding its interpretability problems, in randomized trials. Technical report.

Schmidli, H., Roger, J. H., and Akacha, M. (2021). Estimands for recurrent event endpoints in the presence of a terminal event. *Statistics in Biopharmaceutical Research*, pages 1–11.

Tsiatis, A. (2006). *Semiparametric theory and missing data*. Springer Verlag.

Van der Laan, M. and Robins, J. (2003). *Unified methods for censored longitudinal data and causality*. Springer Verlag.