Modeling time-varying exposure using inverse probability of treatment weights

Nathalie Grafféo*[,1,2], Aurélien Latouche[3,4], Ronald B Geskus[5,6,7], and Sylvie Chevret[1,2,8].

[1]INSERM U1153, Statistic and Epidemiologic Research Center Sorbonne Paris Cité (CRESS), ECSTRA Team, Saint-Louis Hospital, 75010 Paris, France
[2]Paris Diderot University, Paris 7, SPC, Paris, France
[3]Cnam, EA4629, 2 rue conté, 75003 Paris, France
[4]Institut Curie, Inserm U900, Saint Cloud, France
[5]Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Center and Public Health Service Amsterdam, Amsterdam, The Netherlands
[6]Oxford University Clinical Research Unit, Wellcome Trust Major Overseas Programme, Ho Chi Minh City, Viet Nam
[7]Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom
[8]SBIM, Saint-Louis Hospital, APHP, Paris, France

Abstract

For estimating the causal effect of treatment exposure on the occurrence of adverse events, inverse probability weights (IPW) can be used in marginal structural models to correct for time-dependent confounding. The `R` package `ipw` allows IPW estimation by modeling the relationship between the exposure and the confounders via several regression models, among which the Cox model. For right censored data and time-dependent exposures such as treatment switches, the `ipw` package allows a single switch, assuming that patients are treated once and for all. However, to accommodate multiple switches, we extend this package by implementing a function that allows for multiple and intermittent exposure status in the estimation of IPW using a survival model. This extension allows for the whole exposure treatment trajectory in the estimation of IPW. The impact of the estimated weights on the estimated causal effect, with both methods, is assessed in a simulation study. Then, the function is illustrated on a real dataset from a nationwide prospective observational cohort including patients with inflammatory bowel disease. In this study, patients received one or multiple medications (thiopurines, methotrexate and anti-TNF) over time. We used a Cox marginal structural model to assess the effect of thiopurines exposure on the cause-specific hazard for cancer incidence considering other treatments as confounding factors. To this end, we used our extended function which is available online in the Supporting Information.

keywords: Causal inference; Inverse probability weighting; Marginal structural models; `R package`; treatment Switch.

*Corresponding author: e-mail: nathalie.graffeo@univ–amu.fr, Phone: $+33-142-499-742$, Fax: $+33-142-499-745$

1 Introduction

Patients with chronic diseases are prone to many clinical and biological events that may require treatment decisions. To evaluate treatment for such patients in the clinical trial setting the primary endpoint is dedicated to efficacy, while only secondary or tertiary endpoints consider the occurrence of adverse events. Therefore, adverse events reported in clinical trials are those with a high prevalence or a short onset time. Thus it is usually necessary to study large patient cohorts with a long-term follow-up outside any intervention trials to detect rare, and/or possibly delayed outcomes related to a long history of treatment exposure. In this framework, any inference regarding the causal effect of one drug on the occurrence of a particular endpoint is further complicated by the dynamic treatment regimens (as defined in the introduction of Chakraborty and Moodie (2013)). Indeed, administration or discontinuation of the drug of interest may rely on the patient status, confounding the relationships between that drug and the endpoint. Moreover, when the main interest focuses on estimating treatment side effects (rather than efficacy), the history of other administered drugs may act as time-varying confounders, given their potential association with both the decision to prescribe or interrupt the drug of interest and the outcome. Thus, inference on the causal relationship between the treatment and an adverse event is complicated by such intermittent multiple exposures. Those time-dependent treatment decisions act as time-varying confounders, affecting both the occurrence of the event of interest and the patient's exposure to the treatment of interest. Conversely, time-varying confounders are affected by previous exposure to the treatment of interest. For example, health status, medical history, and concomitant therapies can be considered as time-varying confounders acting as intermediate variables at the same time. More precisely, a confounder that is by definition unbalanced between exposed and unexposed patients, can lead to the false conclusion that treatment exposure is associated with adverse events.

In the causal inference framework, it is acknowledged that standard statistical methods may yield a hazard ratio that cannot be interpreted as the true causal effect of the exposure of interest. It is the case when fitting a Cox model with adjustment for time-dependent confounders (for instance, other treatment exposures) because exposure of interest is possibly strongly influenced by such confounding exposures. One approach to handle confounding is to use marginal structural models (Hernán et al., 2000, 2001; Robins et al., 2000; Bailly et al., 2015), in which the parameter estimates of the model can be obtained by Inverse Probability Weights (IPW). More precisely, each observation is weighted by the inverse of the patient's probability of the received treatment given his observed values for the confounders. This creates a pseudo-population in which the exposure is independent of the measured confounders (Cole and Hernán, 2008).

Our interest is to estimate the effect of a treatment of interest on the occurrence of cancer from observational data. Given the competing risks framework induced by deaths free of cancer, a proportional cause-specific hazards Cox model can be used as the marginal structural model that requires estimating the individual weights of being treated in the presence of confounders. Recently van der Wal and Geskus (2011) proposed an `R` package, called `ipw` that allows

handling time-fixed or time-varying exposures of interest, as well as time-fixed or time-varying confounders. However, the weights should be computed according to how the exposure of interest applies. If the exposure is applied over time in some but not all patients, a survival model can be used to model the time to exposure, allowing to handle the right censoring of the exposure. Such a survival model for treatment exposure can be fitted using the `ipw` package proposed by van der Wal and Geskus (2011). Nevertheless, if the exposure of interest is discontinued, then reintroduced, such a complex trajectory cannot be handled in the package. Therefore, our objective is to extend the `ipw` package by proposing and implementing a function that allows time-varying exposure in the computation of weights when the exposure is a right censored endpoint.

The paper is organized as follows. Section 2 presents the CESAME study, the motivating example that aimed at assessing the causal effect of thiopurines exposure on cancer incidence in the presence of time-varying confounders due to other treatments using data from an observational cohort. In Section 3, after giving the formal definition of the IPW, we describe how they can be estimated using the modified `R ipw` package. We illustrate both the existing and the extended functions on a simulated example and compare the impact between the old and new weights on the estimation of the causal effect through simulations. Section 4 provides an application to the CESAME data. Finally, we conclude with a brief discussion.

2 Motivating example

The CESAME cohort is a nationwide prospective observational cohort in which patients with Inflammatory Bowel Disease (IBD) were included. There is no cure for this chronic disease which progresses over time requiring the use of anti-inflammatory drugs and immunosuppressors. This cohort was designed to assess the effect of immunosuppressor exposure, especially that of thiopurines, on cancer incidence in IBD patients. Indeed, these immunosuppressors have been reported to facilitate the genesis of malignancies (Gutierrez-Dalmau and Campistol, 2007). Patients were at risk of developing cancer or to die free of cancer, defining a competing-risks setting. Thus, to estimate the effect of thiopurines exposure on cancer incidence, time-dependent Cox models for the cause-specific hazard of cancer were used in the primary analyses of these data. The primary analysis of the CESAME cohort based on 19,486 patients first showed an excess of risk of lymphoproliferative disorders in patients receiving thiopurines, as compared to those who never received this drug (Beaugerie et al., 2009), with a cause-specific hazard ratio (CSHR) adjusted for age, sex, and duration of IBD, CSHR=5.28 (95% CI= 2.01–13.9). Another analysis of this cohort failed to show any modified risk of colorectal cancer (Beaugerie et al., 2013). However, based on the adjusted time-dependent Cox models, these primary analyses did not account for any time-fixed or time-varying confounders.

Actually, besides thiopurines, two other immunosuppressive drugs were administered to patients from the CESAME cohort, namely methotrexate and anti-Tumour Necrosis Factor (antiTNF), that may act as confounders in the relationship between thiopurines exposure and cancer occurrence as shown in Figure 1. After eliminating patients with incoherent data and patients without

follow-up, we used in the further analyses the 9,075 patients from the cohort previously published (Beaugerie et al., 2009) who were free of these drugs before study enrolment, from May 2004 until June 2005. The end of follow-up was set to December 2007. There were 144 occurrences of cancer and 69 deaths, including 29 caused by cancer. A total of 2,082 patients were exposed at least once to thiopurines (*vs.* 259 for methotrexate and 531 for antiTNF). Among these 2,082 patients who were exposed to thiopurines, there were 24 occurrences of cancer and 5 deaths free of cancer. Note that given the observational nature of the study, observations were made at irregular time intervals.

A proper alternative in a causal inference framework is to fit a Cox marginal structural model with IPW to estimate the causal effect of the dynamic thiopurines regimen on the cause-specific hazard of cancer in the presence of time-varying confounding due to other immunosuppressive treatments. Given that thiopurines exposure may occur irregularly over time, we used a survival model to compute the weights conditionally on the confounders. More precisely, thiopurines exposure can be considered as a recurrent time-to-event outcome. Accordingly, the relationship between thiopurines exposure and the confounders was modeled using the Andersen-Gill model (Andersen and Gill, 1982). Thus, when assessing the relationship between thiopurines exposure and these confounding treatments with a survival model in `R`, we had to extend the `ipwtm` function of the `ipw` package which ignores changes in the exposure of interest status after the start of exposure.

3 Method and simulated example

In this section, after formally defining the IPW, we describe the `R` package `ipw` which computes IPW under several models of the relationship between the treatment exposure and the confounders. We also provide an example of simulated data illustrating both the existing and the extended functions. Source code to reproduce the results is available as Supporting Information on the journal's website (`http:/onlinelibarrywileycom/doi/xxx/suppinfo`).

3.1 IPW definition

Using IPW allows adjusting for confounding by creating a pseudo-population in which the exposure is independent of the measured confounders. Then, for a given patient i, the weight up to the ordered failure time k is defined as:

$$w_{ik} = \prod_{t=1}^{k} \frac{1}{P(A_t = a_{t,i} | \overline{A}_{t-1} = \overline{a}_{t-1,i}, \overline{C}_t = \overline{c}_{t,i}, V = v_i)}, \tag{1}$$

where a_t denotes the treatment that the patient received at time t, $\overline{a}_{t-1}$ denotes the observed treatment exposure history up to time $t-1$, $\overline{c}_t$ denotes the observed history of the time-varying confounders, and v_i denotes the observed time-fixed covariates. Unstable weights can occur when few values of treatment exposure are found in some stratum as defined by the conditioned part of the denominator of the Formula 1. Thus, stabilized weights were suggested by Robins et al.

(2000):

$$sw_{ik} = \prod_{t=1}^{k} \frac{P(A_t = a_{t,i} | \overline{A}_{t-1} = \overline{a}_{t-1,i}, V = v_i)}{P(A_t = a_{t,i} | \overline{A}_{t-1} = \overline{a}_{t-1,i}, \overline{C}_t = \overline{c}_{t,i}, V = v_i)}.$$

Resulting narrower distribution of the weights obtained by stabilization increases the statistical efficiency by producing smaller confidence intervals (van der Wal and Geskus, 2011).

To estimate the causal effect of A_t on the exposure of interest, one can fit a marginal structural model to the observations made at time points t_{ik} weighted by sw_{ik}. Four assumptions underlie the use of marginal structural models in the estimation of the average treatment effect (Cole and Hernán, 2008). The first assumption is consistency, that is, the observed outcome for each patient is the counterfactual outcome that results from each patient's set of observed characteristics. The second assumption is exchangeability, i.e., the absence of unmeasured confounding. The third assumption is positivity that requires the existence of treated and untreated patients at every level of the confounders. Fourth, the model used to estimate the weights has to be well specified.

When the time to treatment defines the exposure of interest, the probability of being treated at time t over an infinitesimal interval $[t, t + \Delta t[$ can be derived from the cumulative hazard of exposure at that time defined by

$$1 - \frac{S(t + \Delta t)}{S(t)} = 1 - \exp\left(- \int_t^{t+\Delta t} \lambda_0(u) \exp(\boldsymbol{\beta}' \boldsymbol{X}(u)) du \right), \qquad (2)$$

where λ_0 stands for the baseline hazard function, $\boldsymbol{X}$ is a vector of covariates and $\boldsymbol{\beta}$ is the vector of regression coefficients. The weights must be estimated at each observed treatment switch in the whole sample, thus computed at each modification of the treatment status of all patients. With a time to event outcome such as the occurrence of cancer, these weights also need to be updated at each failure time from cancer observed in the whole sample.

3.2 Implementing IPW in R: the ipw package

The `ipw` package allows handling time-fixed or time-varying exposures of interest, as well as time-fixed or time-varying confounders (van der Wal and Geskus, 2011). To model the relationship between the treatment exposure and the confounders, the `ipw` package provides five families of regression models which are typical for binomial, Gaussian, multinomial, ordinal, and survival outcomes. Once the link function family is chosen, the weights can be estimated up to the first switch of the treatment of interest, or at all time points. The former option supports the binomial, multinomial, ordinal and survival families, while the latter is only supporting the binomial and Gaussian families.

We created an example dataset (`dataExample`) for illustration. It was generated to mimic a cohort with measures of a time-varying binary treatment exposure of interest (`expo`), a time-varying confounding treatment (`xt`) and a time-fixed binary covariate (`x`). For each patient (identified by `id`), the entire follow-up is divided into several intervals [`tstart`, `fuptime`], each resulting in a separate row, according to the observed changes in either `expo` or `xt` values. Because we are interested here in the weights estimation, the outcome is the

treatment allocation, i.e., `expo`. Then, changes in `expo` are reported at the end of the interval while that on `xt` occur at the onset of the interval. Thus, the variable `expo` takes the value 1 when the treatment is prescribed at the end of the interval (`fuptime`) and 0 otherwise, while `xt` equals 1 when the patient has been exposed to the confounding variable between `tstart` and `fuptime`, and 0 otherwise. Finally, the occurrence of a failure time for each patient is coded 1 whereas a censored observation is coded 0. The observations from `dataExample` of patients `id = 3` and 47 are displayed below.

```
> dataExample[dataExample$id %in% c(3,47), ]
 id   tstart   fuptime   expo   x   xt   event
  3    0.000    0.365      1    1   0      0
  3    0.365    0.417      0    1   0      0
  3    0.417    0.964      0    1   1      0
  3    0.964    1.000      0    1   0      1
 47    0.000    0.340      0    0   0      0
 47    0.340    0.872      1    0   1      0
 47    0.872    0.946      1    0   1      0
 47    0.946    1.000      0    0   1      0
```

The patient with `id = 3` is followed until time 1, which is a failure time. This patient is administered the exposure treatment at time 0.365. Meanwhile, this patient is exposed to the confounding treatment from time 0.417 to time 0.964. Finally, the time-fixed covariate for this patient is set to 1. Conversely, the patient `id = 47` is censored at time 1. He is administered the exposure treatment at time 0.872 and time 0.946, respectively. Meanwhile, he is exposed to the confounding treatment from time 0.340 until the end of his follow-up. Finally, the time-fixed covariate for this patient is set to 0.

In the setting of irregular time intervals, further data processing is required. For the estimation of the likelihood of a proportional hazard model, the individual weights must be available at each failure time in the dataset. Thus, the individual weights are to be updated whenever they change, that is when an individual's exposure of interest or confounders are modified. We defined a data frame `startstop` containing all time points, as defined above, up to each patient's end time as coded in the `ipw` package (van der Wal and Geskus, 2011). However, a new row is merged at each time point where those variables changed value. The stabilized IPW are obtained with the following command:

```
> example.ipw ← ipwtm(  exposure = expo, family = "survival",
                        numerator = ~x, denominator = ~xt + x,
                        id = id, tstart = tstart, timevar = fuptime,
                        type = "first", data = startstop, trunc = 0.01 )
```

that takes the following arguments:

- `exposure` is a vector representing the exposure of interest.
- `family` is used to specify the "survival" family of link functions, used to model the relationship between the variables in `numerator` or `numerator` and `exposure` respectively.
- `numerator` is a formula, specifying the right-hand side of the model used to estimate the elements in the numerator of the IPW. Note that when left unspecified, unstabilized weights with a numerator of 1 are estimated.

- `denominator` is a formula, specifying the right-hand side of the model used to estimate the elements in the denominator of the IPW.
- `id` is a vector, uniquely identifying the patients under observation within which the longitudinal measurements are taken.
- `tstart` is a numerical vector, representing the starting time of follow-up intervals, using the counting process notation.
- `timevar` is a numerical vector, representing follow-up time. This variable is used as the end time of follow-up intervals, using the counting process notation, because `family = ''survival''`.
- `type` specifies the type of exposure. In the `ipw` package, only `''first''` is implemented for the `''survival''` family, meaning that weights are estimated up to the first switch. After this switch, weights will then be constant.
- `data` is a dataframe containing `exposure`, the variables used in `numerator` and `denominator`, and variables `id`, `tstart` and `timevar`.
- `trunc` is an optional fraction for the weights. For instance, when `trunc = 0.01`, the left tail is truncated to the 1st percentile, and the right tail is truncated to the 99th percentile.

The following command allows to find the minimum, the maximum, the mean and the first and ninety-ninth percentiles of the estimated stabilized IPW:

```
> a <- example.ipw$ipw.weights
> c(quantile(a, probs=c(0, 0.01)), mean(a), quantile(a, probs=c(0.99,
1)))
  Min.   1st Percentile.    Mean   99th Percentile    Max.
 0.3540       0.5920       0.9979       2.5298       3.1162
```

Figure 2 shows the distribution of the logarithm of the truncated stabilized IPW obtained by the `ipwtm` function.

3.3 The extended function

Our extended function allows for handling the changes of the time-dependent status. For the sake of simplicity, we assumed that confounding treatments act similarly whatever the rank of the change is. More precisely, we did not consider the whole history of confounding treatments exposure but only considered that the exposure of interest at time t only depended on the confounding treatments given at time t. Then, instead of computing the weights up to the first time when the exposure of interest is modified, the weights are computed for all time points, that is for all changes in the treatment of interest status and for all event times. Following the `ipw` package, the `startstop` coding was employed to format the data frame.

Using the modified function, which has the same syntax as the `ipwtm` function, one may obtain the stabilized and the truncated stabilized weights. Moreover, the modified function produces a list containing, among other things, the inverse probability weights for each observation and the truncated inverse probability weights. Thus, we get the distribution of the former by the following command:

```
> b <- example.ipw.mod$ipw.weights
```

```
> c(quantile(b, probs=c(0, 0.01)), mean(b), quantile(b, probs=c(0.99,
1)))
  Min.   1st Percentile.    Mean   99th Percentile    Max.
 0.1349       0.4553       1.0013       2.4511       11.0541
```

Figure 3 presents the distribution of the logarithm of the truncated stabilized IPW obtained by the extended function. Note that Figure 3, like Figure 2, shows that the mean weights are increasingly negative with time.

The distribution of the relative biases between the theoretical and estimated weights has a median of 0.0002 (interquartile range = -0.006–0.326). Since the simulated data arise from a Weibull baseline hazard with known parameters λ and ν, and known regression coefficients β_x and β_{xt} for both x and xt, the probability of being treated is given by equation 2, that results here in $1 - \exp(-\Lambda_i) = 1 - \exp\{\lambda(t_2^\nu - t_1^\nu)\exp(\beta_x x_i + \beta_{xt} x t_i)\}$, and that of being untreated is given by $\exp(-\Lambda_i)$.

3.4 The impact of the extended function on the estimation of the causal effect

We studied the performance of the extended function on the estimation of the causal effect with a simulation study. We mimicked a cohort of patients who received treatment for some chronic disease; treatment could be discontinued due to the disease status or adverse events. We focused our interest in estimating the treatment effect on the occurrence of adverse events such as cancer, based on a fixed follow-up of 1 year. The sample size was fixed at 500. Recurrent exposure initiation times (`Init`) were generated based on Jahn-Eimermacher et al. (2015) and Austin (2012), depending on the previous value of the treatment exposure, on a time-fixed binary covariate (`x`), and on a time-varying confounding variable (`xt`) at time t and $t-1$. Once initiated, the exposure duration was set at 1 month. `Term` denotes the indicator of the treatment discontinuation. The outcome was generated depending on the two time-dependent binary covariates, i.e., the treatment of interest `expo` and the confounding variable `xt`. The causal hazard ratio of the treatment of interest exposure on the outcome was set at 1.3, and the hazard ratio of the time-varying confounder at 1.2. An excerpt of a simulated dataset looks like:

```
> dataExample[dataExample2$id == 1, ]
 id   tstart   fuptime   Init   Term   x   xt   expo   event
  1    0.000     0.620      1     NA   0    0      0       1
  1    0.620     0.643     NA      0   0    0      1       0
  1    0.643     0.702     NA      0   0    1      1       0
  1    0.702     0.703     NA      1   0    0      1       0
  1    0.703     0.931      1     NA   0    0      0       0
  1    0.931     1.000     NA      0   0    0      1       1
```

For instance, patient with `id=1` is exposed to the treatment of interest twice: (1) from time 0.620 until time 0.703; and (2) from time 0.931 until the end of the follow-up. The `ipwtm` and the extended functions were applied to the events defined by `Init` and `Term`. Second, the weights were obtained by merging

those assessed considering `Init` and `Term` as exposure of interest respectively. Finally, we fitted an Andersen-Gill marginal structural model with `event` as the endpoint, adjusting for `x` and `expo`.

Using the truncated stabilized IPW obtained by the `ipwtm` function, we estimated an absolute bias in the estimation of the causal hazard ratio of `expo` on the time-dependent outcome equal to 0.123 (Empirical Coverage Rate, ECR = 85.4%). Using the weights estimated by the modified function, we obtained a more accurate estimated causal hazard ratio with an absolute bias equal to 0.066 (ECR = 94.8%).

4 Application to CESAME data

The naive analysis was first employed considering a standard Cox model adjusted for both baseline characteristics (`Age` and `Sex`) and time-varying characteristics (treatment of interest exposure and the time-varying treatments confounders) to assess the effect of thiopurines exposure on the cause-specific hazard of cancer. The cause-specific hazard ratio for thiopurines exposure was estimated at 1.01 (95%CI = 0.57–1.78). Then, we applied the extended function to the CESAME data. To take into account the duration of the thiopurines exposure, we considered two events, that is the indicator of thiopurines initiation and that of thiopurines discontinuation. The stabilized weights were computed, using `Age` and `Sex` as baseline fixed covariates.

Then, we fitted Cox marginal structural models with time-varying exposure and confounders, using those stabilized weights and our extended function. We assumed that the structural model was correctly specified because our objective was to provide weights for this particular setting. However, underlying assumption of positivity was checked, with no evidence of structural nor random zeros according to the level of the confounding treatments. The marginal structural model included those variables used to stabilize the weights, namely age (assuming log-linear effect) and gender, to adjust for possible residual confounding as previously recommended (Cole and Hernán, 2008). As highlighted in the Table 1, the stabilized weights have an estimated mean of 1. Thus, the necessary condition for correct model specification was fulfilled.

We found no evidence of any impact of thiopurines exposure on the cause-specific hazard of cancer, whatever the type of cancer with an estimated cause-specific hazard ratio corresponding to the marginal causal effect of thiopurines exposure on cancer incidence of 1.16 (95%CI= 0.67–2.00). The truncation of weights did not markedly modify these estimates. We found similar conclusion from the cause-specific analysis of death free of cancer. The source code is available as Supporting Information on the journal's website.

To sum up, we did not find any causal effect of thiopurines exposure on all type of cancer incidence, while previous papers showed a significant increase in lymphoproliferative disorders (Beaugerie et al., 2009) and a nonsignificant decrease in colorectal neoplasia (Beaugerie et al., 2013). However, there were differences in the handling of data and of statistical modeling that may explain such discrepancies. Firstly, regarding data handling, we excluded the 8,819 pa-

tients who had been treated before inclusion given we didn't know their time of first exposures (for both the treatment of interest and the confounding treatments). Secondly, whereas previous analyses were interested in specific cancers, we considered all types of cancers altogether, notably because there were only 12 lymphomas in the dataset. This assumes that thiopurines exposure acts similarly whatever the site of cancer and this could be unrealistic. Furthermore, the maximal follow-up is around three years, and this cannot be sufficient for a patient to develop cancer. Finally and most importantly, the previous analyses did not handle any potential confounding in the relationships between treatment exposure and cancer incidence, and this is well known to provide biased estimates (see, e.g., Robins et al. (2000)).

5 Discussion

We extended the `ipw` package to take into account multiple and intermittent statuses for the exposure of interest in the estimation of IPW. This is particularly important when dealing with dynamic treatment regimens, that is, in situations where treatments are not taken once and for all. To deal with the presence of extreme weights, we also implemented an option to estimate truncated weights, as it was originally implemented in the `ipw` package. We first focused on a comparison between the weights estimated by our function and the exact weights computed using a generated dataset. This comparison showed that the median of the relative biases was small (0.002). Then, we compared the estimated causal treatment effect obtained using the IPW from the built-in function of the `ipw` package and those obtained by the extended function. Results clearly showed that this new function provided more accurate estimation of the causal treatment effect. A limitation of the present study is that we considered a single value of the hazard ratio of the treatment exposure. Also, we limited the number of treatment switches notably because the data processing in the `startstop` format leads to a huge data frame with millions of rows even with moderate sample sizes. We considered treatment exposure as a binary variable over time in the computation of weights, and we are investigating at this time how to better integrate the history of exposure of interest in these weights. This modeling requires the knowledge of a clinician to mimic efficiently the cumulative treatment effect over time. For example, a naive indicator of current exposure would assume that the treatment effect does not accumulate over time, which is not always the case (Csajka and Verotta, 2006). Then flexible methods allowing to take into account the cumulative treatment effects over time could be of interest (Sylvestre and Abrahamowicz, 2009; Xiao et al., 2014). Otherwise, we assumed that confounding treatments act similarly whatever the rank of the changes in the exposure of interest is. We did not take into account confounders history. This could be addressed by using a different model for each switch in the relationship between the exposure of interest and the confounders. Further work is needed to study this modeling options. The simplest way would be to include a counter of the previous exposure as a covariate in the treatment model.

Regarding the exchangeability assumption, we applied the weights to a Cox marginal structural model assuming that there were no unobserved confounders. Obviously, checking of such assumption from observed data is impossible and is

often based on the knowledge of the confounders in the particular setting. Based on our cohort data, it is notably evident that some of the major confounders, such as tobacco status, were not available. Also, we assumed that consistency held and checked the absence of random zeros (positivity assumption, results not shown). The parametric specification of the treatment model and the model used to estimate the weights should be assessed, including log-linearity and proportional hazards assumptions. Therefore, results from the CESAME cohort should be interpreted with caution. It was mostly used as an illustrative example for time-dependent weights computation and the need for an extension of the `ipw` package. Besides, we were in a competing risk setting, because patients from our dataset only developed cancer once. Using the Andersen–Gill model, our simulation study showed that applications to recurrent events such as adverse events are straightforward.

Acknowledgements
The authors thank the two reviewers and the Associate Editor for all their constructive and helpful comments that have clearly improved our manuscript. We also thank Pr. Laurent Beaugerie for the access to the CESAME data. This work was supported by the *Agence nationale de sécurité du médicament et des produits de santé* (ANSM 2014-S035).

Conflict of Interest
The authors have declared no conflict of interest.

References

Andersen, P. K. and Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. The Annals of Statistics, 10:1100–1120.

Austin, P. C. (2012). Generating survival times to simulate Cox proportional hazards models with time-varying covariates. Statistics in medicine, 31(29):3946–3958.

Bailly, S., Pirracchio, R., and Timsit, J. (2015). What's new in the quantification of causal effects from longitudinal cohort studies: a brief introduction to marginal structural models for intensivists. Intensive care medicine, 42(4):576–579.

Beaugerie, L., Brousse, N., Bouvier, A. M., Colombel, J. F., Lémann, M., Cosnes, J., Hébuterne, X., Cortot, A., Bouhnik, Y., Gendre, J. P., et al. (2009). Lymphoproliferative disorders in patients receiving thiopurines for inflammatory bowel disease: a prospective observational cohort study. The Lancet, 374(9701):1617–1625.

Beaugerie, L., Svrcek, M., Seksik, P., Bouvier, A.-M., Simon, T., Allez, M., Brixi, H., Gornet, J.-M., Altwegg, R., Beau, P., et al. (2013). Risk of colorectal high-grade dysplasia and cancer in a prospective observational cohort of patients with inflammatory bowel disease. Gastroenterology, 145(1):166–175.

Chakraborty, B. and Moodie, E. (2013). Statistical methods for dynamic treatment regimes. Springer.

Cole, S. R. and Hernán, M. A. (2008). Constructing inverse probability

weights for marginal structural models. American Journal of Epidemiology, 168(6):656–664.

Csajka, C. and Verotta, D. (2006). Pharmacokinetic–pharmacodynamic modelling: history and perspectives. Journal of pharmacokinetics and pharmacodynamics, 33(3):227–279.

Gutierrez-Dalmau, A. and Campistol, J. M. (2007). Immunosuppressive therapy and malignancy in organ transplant recipients. Drugs, 67(8):1167–1198.

Hernán, M. Á., Brumback, B., and Robins, J. M. (2000). Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. 11:561–570.

Hernán, M. A., Brumback, B., and Robins, J. M. (2001). Marginal structural models to estimate the joint causal effect of nonrandomized treatments. Journal of the American Statistical Association, 96(454):440–448.

Jahn-Eimermacher, A., Ingel, K., Ozga, A.-K., Preussler, S., and Binder, H. (2015). Simulating recurrent event data with hazard functions defined on a total time scale. BMC medical research methodology, 15(1):16.

Robins, J. M., Hernan, M. A., and Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. Epidemiology, 11:550–560.

Sylvestre, M.-P. and Abrahamowicz, M. (2009). Flexible modeling of the cumulative effects of time-dependent exposures on the hazard. Statistics in medicine, 28(27):3437–3453.

van der Wal, W. M. and Geskus, R. B. (2011). Ipw: an R package for inverse probability weighting. Journal of Statistical Software, 43(13):1–23.

Xiao, Y., Abrahamowicz, M., Moodie, E. E., Weber, R., and Young, J. (2014). Flexible marginal structural models for estimating the cumulative effect of a time-dependent treatment on the hazard: reassessing the cardiovascular risks of didanosine treatment in the swiss HIV cohort study. Journal of the American Statistical Association, 109(506):455–464.

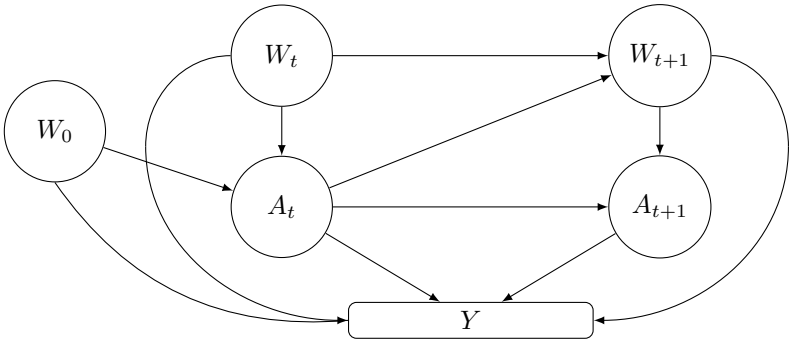

Figure 1: Relationship between the outcome (Y, occurrence of cancer), the treatment of interest (A_t, thiopurines exposure at time t) and the confounding treatments (W_t, other treatments exposure: methotrexate and antiTNF).

Table 1: Cause-specific effect of thiopurines on the occurrence of cancer under progressive truncation of inverse probability weights, CESAME cohort study.

Truncation percentiles	Mean (Min/Max) Estimated stabilized weights		CSHR of cancer*
	relative to treatment initiation	relative to treatment discontinuation	Estimate (95% CI)
0,100	1.000 (0.115–7.827)	1.040 (0.256–44.93)	1.156 (0.669–1.999)
1,99	0.996 (0.958–1.166)	1.023 (0.864–2.617)	1.181 (0.685–2.036)
5,95	0.993 (0.972–1.002)	0.991 (0.924–1.146)	1.184 (0.696–2.017)

* Cause-specific hazard ratio (CSHR) adjusted on gender and age.

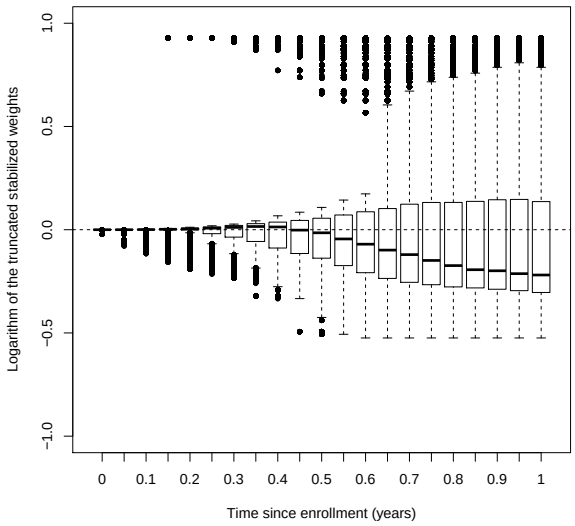

Figure 2: Truncated stabilized weights for the example obtained with the `ipwtm` function that is, when the weights are computed up to the first change in the exposure of interest status. The logarithmic transformation was applied for display purposes.

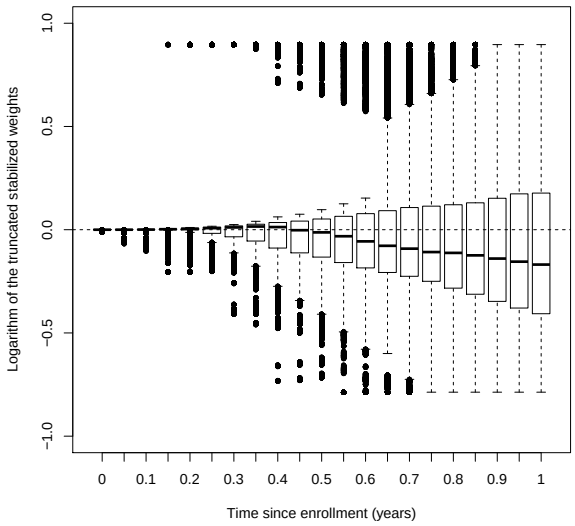

Figure 3: Truncated stabilized weights for the example obtained with the modified function that is, when the weights are computed up for all changes in the exposure of interest status. The logarithmic transformation was applied for display purposes.