A note on the validity of the case-time-control design for autocorrelated exposure histories

Aksel K.G. Jensen
Thomas A. Gerds
Peter Weeke
Christian Torp-Pedersen
Per K. Andersen

Research Report 12/11
Department of Biostatistics
University of Copenhagen
A note on the validity of the case-time-control design for autocorrelated exposure histories

Aksel K.G. Jensen¹, Thomas A. Gerds¹, Peter Weeke², Christian Torp-Pedersen², and Per K. Andersen¹

¹Department of Biostatistics, University of Copenhagen
²Department of Cardiology, Copenhagen University Hospital, Gentofte, Denmark

December 17, 2012

Abstract

The case-time-control design is an extension of the case-crossover design which handles time trends in the exposure of the general population. As in the case-crossover design, time-invariant confounders are controlled for by the design itself. The basic idea is to compare the exposure status of an individual in one or several reference periods where no event occurred to the exposure status of the same individual in the index period where the event occurred. By comparing case-crossover results in cases to case-crossover results in controls exposure-outcome effect can be estimated by conditional logistic regression. We review the mathematical assumptions underlying the case-time-control design. In addition we examine the sensitivity of the design to deviations from the assumed independence of within-individual exposure history. Results from simulating different scenarios suggest that the design is quite robust to deviations from this model assumption. A side result of our investigation shows that the development in exposure probability over time can be modelled in a flexible way, e.g. using polynomials or regression splines.
1 Introduction

Observational cohort studies play an important role for monitoring safety of drugs that have already been approved for marketing. In Weeke et al.\textsuperscript{1} we recently studied the association of prescription antidepressants and the risk of out-of-hospital cardiac arrest by means of the case-time-control design\textsuperscript{2}. The case-time-control design is an extension of the popular case-crossover design. Resulting from an observational study, the results were under the suspicion of an increased risk of bias (compared to a randomized clinical trial). This prompted us to review the mathematical assumptions underlying the case-time-control design.

The case-crossover design was introduced by Maclure\textsuperscript{3} to control for time constant subject characteristics in observational studies. The principle of the case-crossover design is to use cases as their own controls by assessing exposure not only immediately prior to the case period but also prior to a suitably chosen earlier reference period where the event did not occur. The case-crossover design is useful if exposure varies over time and if the latency period from exposure to event is relatively short\textsuperscript{4}. However, case-crossover results are known to be biased when there are time trends in the general population exposure prevalence,\textsuperscript{5} and more generally when there are within-individual exposure dependencies\textsuperscript{6,7}. Vines & Farrington\textsuperscript{6} pointed out that it is safe to analyse case-crossover studies with exactly one control period when there is pairwise exchangeability, which amounts to assuming that the marginal exposure probability is the same in the case period and in the control period. However, when there are multiple control periods valid results can only be expected under a stronger global exchangeability condition, which practically amounts to assuming that all individual exposures across time are independent.

Weeke et al.\textsuperscript{1} studied also antidepressants that were newly introduced on the market. Thereafter, some of them showed strong general population time trends. This made a case-crossover analysis inappropriate, as the exposure effect would be confounded by the period effect. The case-time-control design\textsuperscript{2} was thus suggested to eliminate general population exposure time trends. It works by including controls from the general population. A number \( M \geq 1 \) of controls are matched to each case and their exposure status is analysed in a number \( K \geq 1 \) of periods that are matched to the periods of the corresponding case. As in the case-crossover design, time-invariant confounders are controlled for by the design itself. In addition, general population time trends can be distinguished from exposure-outcome effects.

The purpose of the present work is threefold. First, we will clarify the assumptions leading to Suissa’s logistic model\textsuperscript{2} which combines a model for exposure time trends in the general population with a model for the association of exposure and the risk of an event. Second, we extend Suissa’s framework in several ways. We discuss the details of the design with two reference periods. Furthermore, our exposure model includes time constant subject effects which are allowed to be different from the time constant subject effects on the event risk. Based on this model we are able to show that certain between-subject differences in exposure time trends will not introduce bias. Another extension of practical and theoretical interest is that the exposure probability does not have to be linear. In fact it can be modelled as a function of time in a flexible way, e.g., by using splines. This means that the placement of reference periods is not constrained to be equidis-
tant, and hence can be adapted to other features of the study, such as the nature of the exposure.

Third, we show results from a simulation study on the sensitivity of the case-time-control design in situations where there besides a general population exposure time trend also are within-subject exposure dependencies in the simulated data. To what extent the latter affect the results of the case-time-control analysis is not obvious. In particular not for multiple reference periods (K > 1) which was not explicitly discussed by Suissa.

For further discussion of the case-time-control design see also Suissa\textsuperscript{8} and Greenland\textsuperscript{9}. Alternative ways of addressing the challenge of a period effect in the framework of case-only studies are studied by Whitaker et al.\textsuperscript{10, 11} and Wang et al.\textsuperscript{12}.

2 The case-time-control design

The study period is a time interval \([a, b]\) defined by calendar dates \(a\) and \(b\). The study population consists of all subjects that are alive and event-free at \(a\). For the case-time-control design the interval is divided into non-overlapping time periods, e.g. by using weeks or months as time units. Cases are all subjects that have experienced the event of interest in \([a, b]\). For each case there is a period \(t_0\), which we call the index period hereafter, in which the event occurred. We use the notation \(C(t_0) = 1\) to say that the event occurred in period \(t_0\). We assume an event can only occur once, that is, \(C(t_0) = 1\) implies an event free past: \(C(u) = 0\), \(u < t_0\). To cover situations where not all cases are studied, we denote by \(q_1\) the probability for sampling a case. Thus, if all cases are studied, then \(q_1 = 1\). Controls are subjects that are event-free at \(b\). In all what follows we consider a matched design, where for each case a number \(M\) of controls are sampled from the controls that are alive at the case’s index period. In this way also controls have an index period. For controls we have \(C(t_0) = 0\). The probability of sampling a control will be denoted by \(q_0\). In what follows we only consider binary exposures and we assume that the exposure status is available for all sampled subjects in the time periods before and including the index period. We use notation \(E(t) = 1\) for exposure in period \(t\) and \(E(t) = 0\) for non-exposure in period \(t\).

2.1 Modelling exposure and event risk

Following Suissa\textsuperscript{2} we introduce a logistic model for the risk of an event:

\[
\text{logit}\{P(C_i(t) = 1 \mid B_i, E_i(t) = e, C(u) = 0, u < t)\} = \beta_0 + B_i + \beta e. \tag{1}
\]

where \(B_i\) is an effect for individual \(i\), \(\beta_0\) an average effect, and \(\beta\) the effect of exposure. The model is defined for all time periods in the study interval \([a, b]\), in which we code time as \(t = 0, 1, \ldots, T\), such that \((t = 0)\) corresponds to \(a\) and \((t = T)\) to \(b\).

We use a second logistic model to describe the probability of exposure:

\[
\text{logit} \{P(E_i(t) = 1 \mid A_i)\} = \alpha_0 + A_i + f(t), \tag{2}
\]

where \(\alpha_0\) is the exposure risk at the time origin, \(A_i\) is an effect for individual \(i\) and \(f(t)\) is the effect for the period.
This extends Suissa’s framework as follows. Our model allows that different subjects have different exposure time trends (same functional form but different intercepts). At the same time, the functional form of the time effect is not constrained to be linear. For example, \( f(t) \) could be a parametrized as \( f(t) = \alpha_1 t + \alpha_2 t^2 \). In practice, it is often reasonable to assume that the time trend is a smooth function of time, in which case \( f \) can be approximated by a spline. Furthermore, in our framework the subject effects \( A_i \) and \( B_i \) on exposure and event risk are allowed to have different values.

### 2.2 One reference period

Suissa\(^2\) introduced the case-time-control design with a single reference period for each subject. In what follows we denote the reference period by \( t_1 \) and assume \( t_1 < t_0 \) for all subjects. An additional “wash-out” period that limits possible carry-over effect can be added between \( t_1 \) and \( t_0 \). There may occur practical problems when \( t_1 < a \), that is when the reference period is before calendar date \( a \). Then, one can either exclude these cases, or if this is available collect exposure information also from time periods before the study interval.

The exposure history of a subject is called discordant if \( E(t_0) + E(t_1) = 1 \). We observe that the probability of a discordant exposure history depends on the parameter of interest, \( \beta \), and on the period effects \( f(t_0) \) and \( f(t_1) \). In particular, we show in the appendix that for a case we have

\[
P(E(t_0) = 1, E(t_1) = 0 \mid \sum_{j=0}^{1} E(t_j) = 1, C(t_0) = 1, S = 1) = \frac{e^{\beta + f(t_0)}}{e^{f(t_1)} + e^{\beta + f(t_0)}} \quad (3)
\]

and for a control

\[
P(E(t_0) = 1, E(t_1) = 0 \mid \sum_{j=0}^{1} E(t_j) = 1, C(t_0) = 0, S = 1) = \frac{e^{f(t_0)}}{e^{f(t_1)} + e^{f(t_0)}} \quad (4)
\]

Here \( S = 1 \) is notation for being sampled, see the beginning of this section were we introduced that \( P(S = 1 \mid C(t_0) = 1) = q_1 \) and \( P(S = 1 \mid C(t_0) = 0) = q_0 \). It is important to note that the right hand sides of equations (3) and (4) do not depend on the sampling probabilities. Note that when we condition on \( C(t_0) \) in (3) and (4) we tacitly assume an event free past: \( C(u) = 0, \ u < t_0 \).

The exposure data from \( i = 1, \ldots, n \) subjects can be divided in three groups. The first group are all subjects with concordant exposure history, that is either \( E(t_1) = 0 \) and \( E(t_0) = 0 \) or \( E(t_1) = 1 \) and \( E(t_0) = 1 \). Since there is no information in these data regarding the parameter of interest they can be removed. The second group consists of subjects who where exposed in the reference period and unexposed in the index period: \( E(t_1) = 1 \) and \( E(t_0) = 0 \). This group will be denoted by \( \mathcal{E}_{10} \). Similarly denote by \( \mathcal{E}_{01} \) the group of subjects that were exposed in the index period and unexposed in the reference period: \( E(t_1) = 0 \) and \( E(t_0) = 1 \).

The parameter \( \beta \) can be estimated by maximizing the following likelihood function:

\[
\prod_{i \in \mathcal{E}_{01}} \text{expit} \{ \beta C_i(t_0) + \Delta f_{01} \} \times \prod_{i \in \mathcal{E}_{10}} \left( 1 - \text{expit} \{ \beta C_i(t_0) + \Delta f_{01} \} \right) \quad (5)
\]
where we use the notation $\Delta f_{01} = f(t_0) - f(t_1)$ for the period effect and define $\expit(x) = e^x/(1 + e^x)$.

To deduce the likelihood (5) one has to assume independence of within-subject exposure history given individual random effects

$$P(E_i(t_0) \mid E_i(t_1), A_i, B_i) = P(E_i(t_0) \mid A_i).$$

saying that given the random effects the probability of an individual being exposed in a certain period is the same whether or not the individual has been exposed in an earlier reference period.

### 2.3 Two reference periods

For two reference periods at $t_2$ and $t_1$, exposure-discordant cases and controls include subjects for whom $0 < E(t_0) + E(t_1) + E(t_2) < 3$. As in the one reference period setup conditional probabilities can be computed in a similar way. For discordant cases e.g., as

$$P(E(t_0) = 1, E(t_1) = 0, E(t_2) = 1 \mid E(t_0) + E(t_1) + E(t_2) = 2, C(t_0) = 1, S = 1)$$

$$= \frac{e^{\beta + f(t_0) + f(t_2)}}{e^{\beta + f(t_0) + f(t_2)} + e^{\beta + f(t_0) + f(t_1)} + e^{f(t_2) + f(t_1)}}. \tag{6}$$

again we tacitly condition on an event free past: $C(u) = 0, u < t_0$.

If we extend the one reference period notation in a natural way to the two reference period setup, the combined likelihood function consisting of cases and controls can be written as

$$\prod_{i \in E_{001}} \left(1 + \exp \left(-\beta C_i(t_0) + \Delta f_{10}\right) + \exp \left(-\beta C_i(t_0) + \Delta f_{20}\right)\right)^{-1}$$

$$\times \prod_{i \in E_{010}} \left(1 + \exp \left(\beta C_i(t_0) + \Delta f_{01}\right) + \exp \left(\Delta f_{21}\right)\right)^{-1}$$

$$\times \prod_{i \in E_{100}} \left(1 + \exp \left(\beta C_i(t_0) + \Delta f_{02}\right) + \exp \left(\Delta f_{12}\right)\right)^{-1}$$

$$\times \prod_{i \in E_{011}} \left(1 + \exp \left(\Delta f_{21}\right) + \exp \left(-\beta C_i(t_0) + \Delta f_{20}\right)\right)^{-1}$$

$$\times \prod_{i \in E_{101}} \left(1 + \exp \left(\Delta f_{12}\right) + \exp \left(-\beta C_i(t_0) + \Delta f_{10}\right)\right)^{-1}$$

$$\times \prod_{i \in E_{110}} \left(1 + \exp \left(\beta C_i(t_0) + \Delta f_{02}\right) + \exp \left(\beta C_i(t_0) + \Delta f_{01}\right)\right)^{-1} \tag{7}$$

where we use the notation $\Delta f_{21} = f(t_2) - f(t_1)$, $\Delta f_{02} = f(t_0) - f(t_2)$ etc. to describe period effects. Computations leading to (6) can be found in the appendix.
As for one reference period one has to assume independence of within-subject exposure history given individual random effects to deduce the likelihood term (7).

\[
P(E_i(t_0) \mid E_i(t_1), E_i(t_2), A_i, B_i) = P(E_i(t_0) \mid A_i).
\]

The simulation study reported in section 3 investigates robustness to deviations from this assumption in 6 different setups analyzed with both one and two reference periods.

### 2.4 Special cases of the case-time-control design

Further assumptions can be made about the exposure time-trend \( f(t) \). For notational simplicity we only consider one reference period below, i.e. special cases of (5) but analogous special cases can be considered for two reference periods.

**The case-crossover design:** One assumes no time-trend i.e., \( f(t) = \text{constant} \). The controls are ignored as they do not contain any information about \( \beta \).

\[
\prod_{i \in E_{01}} \expit(\beta) \times \prod_{i \in E_{10}} \left(1 - \expit(\beta)\right)
\]

**The case-time-control design with linear time-trend:** A linear time-trend in exposure on the logit-scale implies \( f(t) = \alpha_1 t \). The controls are here needed as we can not distinguish the time-trend effect \( \alpha_1 \) from the effect of exposure on outcome \( \beta \) using just the cases.

\[
\prod_{i \in E_{01}} \expit(\beta C_i(t_0) + \alpha_1) \times \prod_{i \in E_{10}} \left(1 - \expit(\beta C_i(t_0) + \alpha_1)\right)
\]

(8)

here the difference between two time periods is scaled to unit size i.e., \( t_0 - t_1 = 1 \).

**The case-time-control design with non-linear time-trend:** One models the time-trend using a polynomial e.g., \( f(t) = \alpha_1 t + \alpha_2 t^2 \). The controls are still needed as the linear part of the time-trend effect \( \alpha_1 \) can not be separated from the effect of exposure on outcome \( \beta \) using just the cases.

\[
\prod_{i \in E_{01}} \expit(\beta C_i(t_0) + \alpha_1 + \alpha_2(t_0^2 - t_1^2)) \times \prod_{i \in E_{10}} \left(1 - \expit(\beta C_i(t_0) + \alpha_1 + \alpha_2(t_0^2 - t_1^2))\right)
\]

(9)

again the difference between two time periods is scaled to unit size and \( t_0 \) and \( t_1 \) are individual specific.

Also any likelihood function (5) where \( f(t) \) is defined via regression splines constitutes a usable special case of the case-time-control design. As mentioned earlier all special cases can be extended to the two reference period setup. The difference is that we actually do not need the controls when we have two reference period as the distinction problems between \( \beta \) and a linear time-trend \( \alpha_1 \) disappear. The main advantage of still using controls in a two reference setup is therefore to gain power and increase robustness against misspecification of \( f \).
2.5 Implementation

Case-time-control analysis can be implemented in SAS (version 9.2, SAS, Inc., Cary, NC) using proc logistic. Data should be organised with one record per person for each period. Exposure-status should be modelled as dependent variable and event-status and time as predictors. Finally the strata=person option ensures that it is the conditional likelihood that is being maximized. Another option is to model exposure status as a survival status at an artificial chosen dummy survival time, using e.g., the phreg procedure in SAS or the coxph function in R (version 2.12.2).

3 Simulations

3.1 Setup

We now describe a simulation study where we examine how deviations from the assumption of independent within-subject exposure history might bias the exposure-outcome effect.

Our aim is partly to reflect the recent study of the risk of out-of-hospital cardiac arrest and use of antidepressants by Weeke et al.\textsuperscript{1} so we simulate scenarios similar that study. Thus, individuals are observed in a 5 year period, with one discrete variable time with values 1,2,...,60, equivalent to 60 months, \textit{time} = 1 reflecting the end of the first month.

Some of our simulations reflect autocorrelated exposure histories close to the ones estimated in Weeke et al.\textsuperscript{1} while others have minor or no autocorrelated exposure histories.

In accordance with (2) but now with an added autocorrelation term $\kappa$, we model the exposure probability for individual $i$ at time $t$ given exposure in the previous period, as

$$\expit\left(\alpha_0 + A_i + f(t) + \kappa \times E_i(t-1)\right)$$

where $\alpha_0$ is baseline exposure-risk parameter and $A_i \sim N(0, \sigma)$ is a random normally distributed variable drawn for each individual reflecting the individual exposure-risk. Further in accordance with Weeke et al.\textsuperscript{1} we choose to model a linear time-trend: $f(t) = \gamma t$. Note that a high value of $\kappa$ implies a high probability that an exposed individual remains exposed in the following period in contradiction with the the assumption of independent within-subject exposure history.

Following (1) we model the probability of event for individual $i$ at time $t$ given exposure history by

$$\expit\left(\beta_0 + B_i + \beta E_i(t)\right)$$

where $\beta_0$ is a baseline event risk parameter and $B_i \sim N(0, \xi)$ is a random normally distributed variable drawn for each individual reflecting the individual event risk. The parameter of main interest $\beta$ is the effect of exposure on event risk.

Strong autocorrelation $\kappa$ will increase the proportion of individuals exposed as time increases as exposed individuals then tend to remain exposed. High values of the individual exposure-risk deviation $\sigma$ amplifies this phenomenon. These increasing exposure tendencies can though be corrected
by a suitably chosen value of $\alpha_0$ so that the overall proportion of individuals exposed does not increase towards one as time increases. For each choice of $\kappa$ and $\sigma$ in the simulations we therefore choose an $\alpha_0$ using pre-simulations (not reported) such that in a setup with no time-effect i.e., $\gamma = 0$, 5\% of the population will be exposed both at time $= 1$ and $time = 60$.

We let $\gamma$ take the fixed value of $\log(1.03)$ in all reported simulations. The proportion exposed at $time = 60$ will then differ between 7\% and 18\% (not reported) depending on the choice of $\kappa$, $\sigma$ and $\alpha_0$ in the six main simulation scenarios reported in Table 1. The effect of exposure $\beta$ in these simulations is kept fixed at $\log(1.40)$, but additional simulations are also reported in Table 2 where the true exposure-outcome effect is smaller or even non-existing ($\beta = 0$).

A naive OR-estimation of $\kappa$ in the data from Weeke et al.\textsuperscript{1} suggests a $\kappa$ value close to 9. This value is found comparing odds for exposure at time $t_1$ given exposure at time $t_2$ with odds for exposure at time $t_1$ given non-exposure at time $t_2$. In short

$$\hat{\kappa}_{\text{naive}} = \log \left( \frac{N_{11}}{N_{01}} \right) \frac{N_{10}}{N_{00}}$$

where, e.g $N_{11}$ is the number of subjects exposed in both reference periods and $N_{01}$ is the number of controls non-exposed in the second reference period ($t_2$) and exposed in the first reference period ($t_1$).

This naive $\kappa$-estimate, however, is biased upwards as it also reflects, in part, the autocorrelation induced by the random effects $A_i$. Again one should notice that high values of $\kappa$ and the exposure-risk standard deviation $\sigma$ lead to an increasing proportion of exposed individuals as time increases, regardless of exposure time-trend. In our simulated data we therefore choose different combinations of $\kappa$ and $\sigma$ which in the simulated data lead to a naive estimate of $\kappa$ close to 9.

Finally we choose fixed values of the event risk baseline parameter $\beta_0$ equal to $-6$ and the personal event risk deviation $\xi = 0.5$ in all our simulations. In two of the three models investigated in the simulated data we introduce a wash-out period of length 30 days between period 0 and period 1. The wash-out period is introduced to reduce a possible carry-over effect if an individual has been exposed in the first reference period.

### 3.2 Results

The main simulation results are found in Table 1, where 6 different parameter combinations are reported. The first three combinations have a $\sigma$-value equal to 2 and show how a $\kappa$-value equal to 0, meaning no within-subject dependent exposure history, implies consistent estimates for all three designs. Further, one sees how increasing values of $\kappa$ lead to increasing bias in $\beta$ for the design with two reference periods but not for the designs with one reference period. The last parameter-combinations aim to reflect possible autocorrelation structures in the data from Weeke et al.\textsuperscript{1} with a $\hat{\kappa}_{\text{naive}}$ value close to 9.
Autocorrelation induced by random effects causes no bias in $\beta$. In contrast, autocorrelation induced by increasing value of $\kappa$ leads to moderate bias in the design with two reference periods and for high $\kappa$-values even a minor bias for the designs with only one reference period. In addition, the ability for all three models to estimate the period effect $\gamma$ correctly is clearly reduced with increasing $\kappa$ values.

Overall, the main advantage of the two reference period design is the power gained. This is most clearly seen by the values of $\text{sd}(\hat{\beta})$ in the simulations and the median percentages of discordant cases contributing to the analysis.

A natural concern when using variants of the case-time-control design is whether the design itself can induce a artificial effect of exposure on outcome when large autocorrelation is present in the data analysed. To further explore this concern we analyse simulated data with considerable autocorrelation and random effect present and with no or only a minor effect of exposure on outcome. The results are found in Table 2 where we investigate the parameter combination $\kappa = 6$ and $\sigma = 4$ which seems to reflect the data from Weeke et al.\textsuperscript{1} best.

We only see a minor bias for $\hat{\beta}$ in the two reference period design which even decreases as the true $\beta$ decreases towards 0. The one reference period designs do not seem affected by the high autocorrelation at all in estimating $\beta$. Overall all three designs severely overestimate the time-effect $\gamma$.


### 3.3 Discussion

Our simulations suggest that a possible bias in the exposure-outcome effect can be introduced in the case-time-control design by choosing several reference periods if there is a strong non-random induced autocorrelation present in the data. This bias seems minor in our simulations, especially when the true effect of exposure on outcome decreases as seen in Table 2. In contrast, designs with only one reference period seem very robust to deviations from the model assumption of independent within-subject exposure history in all our simulations.

The good performance of the one reference period designs could be explained partly by fulfilment of the exchangeability assumption, which in this setup seems quite reasonable. Think e.g., of exposures as being stationary, then the assumption $P(E_i(t_1) = 0, E_i(t_0) = 1) = P(E_i(t_1) = 1, E_i(t_0) = 0)$ only asserts that conditioned on seeing a change in exposure, the probability that it is a person who starts e.g., a treatment is the same as the probability that it is a person who ends a treatment. Assuming these probabilities are the same, ignoring the period effect in this specific term apparently only induces a minor bias. It should be emphasized that the period effect is still accounted for in our model and the actual likelihood (5). It is just in the assumption of exchangeability that one ignores the period effect.

A weakness in the simulations is the reduced ability for all three models to estimate the period...
effect correctly if the non-random autocorrelation increases. This weakness does not seem to confound our effect-estimates $\hat{\beta}$ severely in these simulations but might do in other setups.

Although promising for the general applicability of the case-time-control design in scenarios with some dependent within-subject exposure history one should be cautious using the design especially with several reference periods. Pre-assessment of possible dependent within-subject exposure history should be considered and if possible simulation studies reflecting the specific correlation structure performed. Unless there is a strong need for extra power analysis with one reference period should be preferred compared to analysis with several reference periods. In the latter case a comparison with results from a one reference period analysis is recommended. If the results differ much a possible bias due to autocorrelated exposure-histories in the analysis with several reference periods should be suspected.

Generally, increasing the distance between reference periods should decrease bias from autocorrelated exposure histories, at the cost of introducing possible time-varying confounding. Depending on the understanding of underlying biological mechanism changing the length and placement of periods are possibilities which could be considered as part of a sensitivity analysis too. How information and results from such different versions of a case-time-control design should be used and interpreted and to what extent it helps identify bias caused by autocorrelated exposure histories are though open questions which are beyond the scope of this paper.
TABLE 1. Averages of $\hat{\beta}$, $\hat{\gamma}$, sd and $\hat{\kappa}_{\text{naive}}$ based on 1000 simulations.\textsuperscript{a}

<table>
<thead>
<tr>
<th>Model $^b$</th>
<th>$\exp(\hat{\beta})$</th>
<th>sd($\hat{\beta}$)</th>
<th>Abs.bias</th>
<th>MSE</th>
<th>$\exp(\hat{\gamma})$</th>
<th>sd($\hat{\gamma}$)</th>
<th>$\hat{\kappa}_{\text{naive}}$</th>
<th>Discordant$^d$</th>
<th>$\kappa$</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2R+W</td>
<td>1.400</td>
<td>0.042</td>
<td>0.034</td>
<td>0.002</td>
<td>1.030</td>
<td>0.008</td>
<td>2.04</td>
<td>3.1%</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>1R+W</td>
<td>1.400</td>
<td>0.050</td>
<td>0.042</td>
<td>0.003</td>
<td>1.030</td>
<td>0.012</td>
<td>2.04</td>
<td>2.1%</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>1R</td>
<td>1.400</td>
<td>0.050</td>
<td>0.040</td>
<td>0.002</td>
<td>1.030</td>
<td>0.023</td>
<td>2.04</td>
<td>2.1%</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>2R+W</td>
<td>1.421</td>
<td>0.046</td>
<td>0.039</td>
<td>0.002</td>
<td>1.038</td>
<td>0.008</td>
<td>2.94</td>
<td>2.6%</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1R+W</td>
<td>1.400</td>
<td>0.053</td>
<td>0.041</td>
<td>0.003</td>
<td>1.036</td>
<td>0.012</td>
<td>2.94</td>
<td>1.9%</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1R</td>
<td>1.402</td>
<td>0.057</td>
<td>0.046</td>
<td>0.003</td>
<td>1.042</td>
<td>0.027</td>
<td>2.94</td>
<td>1.6%</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2R+W</td>
<td>1.443</td>
<td>0.050</td>
<td>0.049</td>
<td>0.004</td>
<td>1.050</td>
<td>0.009</td>
<td>3.84</td>
<td>2.2%</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>1R+W</td>
<td>1.402</td>
<td>0.058</td>
<td>0.045</td>
<td>0.003</td>
<td>1.047</td>
<td>0.013</td>
<td>3.84</td>
<td>1.6%</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>1R</td>
<td>1.401</td>
<td>0.066</td>
<td>0.052</td>
<td>0.004</td>
<td>1.061</td>
<td>0.031</td>
<td>3.84</td>
<td>1.2%</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2R+W</td>
<td>1.486</td>
<td>0.157</td>
<td>0.149</td>
<td>0.036</td>
<td>1.215</td>
<td>0.028</td>
<td>9.07</td>
<td>0.2%</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>1R+W</td>
<td>1.419</td>
<td>0.185</td>
<td>0.148</td>
<td>0.035</td>
<td>1.207</td>
<td>0.043</td>
<td>9.07</td>
<td>0.2%</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>1R</td>
<td>1.425</td>
<td>0.247</td>
<td>0.209</td>
<td>0.066</td>
<td>1.406</td>
<td>0.114</td>
<td>9.07</td>
<td>0.1%</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>2R+W</td>
<td>1.477</td>
<td>0.158</td>
<td>0.147</td>
<td>0.034</td>
<td>1.100</td>
<td>0.028</td>
<td>8.72</td>
<td>0.2%</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>1R+W</td>
<td>1.412</td>
<td>0.184</td>
<td>0.145</td>
<td>0.034</td>
<td>1.093</td>
<td>0.043</td>
<td>8.72</td>
<td>0.2%</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>1R</td>
<td>1.398</td>
<td>0.233</td>
<td>0.189</td>
<td>0.057</td>
<td>1.159</td>
<td>0.109</td>
<td>8.72</td>
<td>0.1%</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>2R+W</td>
<td>1.469</td>
<td>0.167</td>
<td>0.151</td>
<td>0.037</td>
<td>1.071</td>
<td>0.030</td>
<td>8.67</td>
<td>0.2%</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>1R+W</td>
<td>1.410</td>
<td>0.192</td>
<td>0.152</td>
<td>0.036</td>
<td>1.064</td>
<td>0.045</td>
<td>8.67</td>
<td>0.1%</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>1R</td>
<td>1.397</td>
<td>0.234</td>
<td>0.187</td>
<td>0.055</td>
<td>1.103</td>
<td>0.110</td>
<td>8.67</td>
<td>0.1%</td>
<td>3</td>
<td>12</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Each simulated population of size N=100000.

\textsuperscript{b} One or two reference periods. Additional wash-out period indicated with +W.

\textsuperscript{c} True $\beta = \log(1.40)$ and true $\gamma = \log(1.03)$

\textsuperscript{d} Median percentages of discordant cases for the 1000 simulations. Median percentages of discordant+concordant cases was 14.3% - 14.6% in the simulations.
TABLE 2. Varying $\beta$ for one fixed parameter combination.$^a$

<table>
<thead>
<tr>
<th>Model</th>
<th>$\exp(\hat{\beta})$</th>
<th>sd($\hat{\beta}$)</th>
<th>Abs.bias</th>
<th>MSE</th>
<th>$\exp(\hat{\gamma})$</th>
<th>sd($\hat{\gamma}$)</th>
<th>$\hat{\kappa}_{\text{naive}}$</th>
<th>$\exp(\beta)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2R+1W</td>
<td>1.005</td>
<td>0.166</td>
<td>0.158</td>
<td>0.038</td>
<td>1.212</td>
<td>0.028</td>
<td>9.08</td>
<td>1.00</td>
</tr>
<tr>
<td>1R+1W</td>
<td>1.098</td>
<td>0.194</td>
<td>0.156</td>
<td>0.038</td>
<td>1.21</td>
<td>0.043</td>
<td>9.08</td>
<td>1.00</td>
</tr>
<tr>
<td>1R</td>
<td>1.003</td>
<td>0.259</td>
<td>0.208</td>
<td>0.067</td>
<td>1.413</td>
<td>0.115</td>
<td>9.08</td>
<td>1.00</td>
</tr>
<tr>
<td>2R+1W</td>
<td>1.11</td>
<td>0.163</td>
<td>0.153</td>
<td>0.037</td>
<td>1.211</td>
<td>0.028</td>
<td>9.07</td>
<td>1.10</td>
</tr>
<tr>
<td>1R+1W</td>
<td>1.095</td>
<td>0.191</td>
<td>0.158</td>
<td>0.039</td>
<td>1.206</td>
<td>0.043</td>
<td>9.07</td>
<td>1.10</td>
</tr>
<tr>
<td>1R</td>
<td>1.095</td>
<td>0.254</td>
<td>0.213</td>
<td>0.072</td>
<td>1.396</td>
<td>0.114</td>
<td>9.07</td>
<td>1.10</td>
</tr>
<tr>
<td>2R+1W</td>
<td>1.229</td>
<td>0.16</td>
<td>0.153</td>
<td>0.036</td>
<td>1.213</td>
<td>0.028</td>
<td>9.07</td>
<td>1.20</td>
</tr>
<tr>
<td>1R+1W</td>
<td>1.201</td>
<td>0.188</td>
<td>0.157</td>
<td>0.038</td>
<td>1.206</td>
<td>0.043</td>
<td>9.07</td>
<td>1.20</td>
</tr>
<tr>
<td>1R</td>
<td>1.206</td>
<td>0.252</td>
<td>0.206</td>
<td>0.068</td>
<td>1.396</td>
<td>0.114</td>
<td>9.07</td>
<td>1.20</td>
</tr>
<tr>
<td>2R+1W</td>
<td>1.355</td>
<td>0.158</td>
<td>0.153</td>
<td>0.037</td>
<td>1.217</td>
<td>0.028</td>
<td>9.07</td>
<td>1.30</td>
</tr>
<tr>
<td>1R+1W</td>
<td>1.306</td>
<td>0.187</td>
<td>0.153</td>
<td>0.036</td>
<td>1.209</td>
<td>0.043</td>
<td>9.07</td>
<td>1.30</td>
</tr>
<tr>
<td>1R</td>
<td>1.308</td>
<td>0.25</td>
<td>0.197</td>
<td>0.061</td>
<td>1.413</td>
<td>0.114</td>
<td>9.07</td>
<td>1.30</td>
</tr>
</tbody>
</table>

$^a$ Averages of $\hat{\beta}$, $\hat{\gamma}$, sd and $\hat{\kappa}_{\text{naive}}$ based on 1000 simulations each of size N=100000, $\kappa = 6$, $\sigma = 4$ and $\exp(\gamma) = 1.03$ in all simulations.
References


Appendix

In the following we define \( \expit(x) = \exp(x)/(1 + \exp(x)) \) and repeatedly use that \( \expit(x)/(1 - \expit(x)) = \exp(x) \). We suppress the person specific index \( i \) in part of the calculations, as we e.g., write \( E(t_0) \) instead of \( E_i(t_{i,0}) \) to denote the exposure status for person \( i \) in the index period belonging to this person. We use the shorthand notation \( S = 1 \) for being sampled to the study, \( B_i \) is an individual specific effect for event and \( A_i \) an individual specific effect for exposure.

Finally in some parts, if needed due to lack of space we will write “case” or “control” in the formulas instead of the more formal \( (C(t_0) = 1, C(u) = 0, u < t_0) \) for a case and \( (C(t_0) = 0, C(u) = 0, u < t_0) \) for a control.

### 3.4 Likelihood-calculations for 1 reference period

#### 3.4.1 Cases:

**Lemma 3.4.1.**

\[
(1) \quad P(E(t_0) = 1, E(t_1) = 0 \mid \sum_{j=0}^{1} E(t_j) = 1, C(t_0) = 1, C(u) = 0, u < t_0, A_i, B_i, S = 1) = \frac{e^{\beta + f(t_0)}}{e^{f(t_1)} + e^{\beta + f(t_0)}}
\]

\[
(2) \quad P(E(t_0) = 0, E(t_1) = 1 \mid \sum_{j=0}^{1} E(t_j) = 1, C(t_0) = 1, C(u) = 0, u < t_0, A_i, B_i, S = 1) = \frac{e^{f(t_1)}}{e^{f(t_1)} + e^{\beta + f(t_0)}}
\]

**Proof.**

\[
P(E(t_0) = 1, E(t_1) = 0 \mid \sum_{j=0}^{1} E(t_j) = 1, C(t_0) = 1, C(u) = 0, u < t_0, A_i, B_i, S = 1)
\]

\[
= \frac{P(E(t_0) = 1, E(t_1) = 0, A_i, B_i, S = 1, \text{case})}{P(E(t_0) = 1, E(t_1) = 0, A_i, B_i, S = 1, \text{case}) + P(E(t_0) = 0, E(t_1) = 1, A_i, B_i, S = 1, \text{case})}
\]

\[
= \left(1 + \frac{P(E(t_0) = 0, E(t_1) = 1, A_i, B_i, S = 1, \text{case})}{P(E(t_0) = 1, E(t_1) = 0, A_i, B_i, S = 1, \text{case})}\right)^{-1}
\]

\[
= \left(1 + \frac{P(S = 1 \mid P(E(t_0) = 0, E(t_1) = 1, A_i, B_i, \text{case}) P(E(t_0) = 0, E(t_1) = 1, A_i, B_i, \text{case})}{P(S = 1 \mid P(E(t_0) = 1, E(t_1) = 0, A_i, B_i, \text{case}) P(E(t_0) = 1, E(t_1) = 0, A_i, B_i, \text{case})}\right)^{-1}
\]

\[
= \left(1 + \frac{q_1 P(E(t_0) = 0, E(t_1) = 1, C(t_0) = 1, C(u) = 0, u < t_0, A_i, B_i) P(E(t_0) = 1, E(t_1) = 0, C(t_0) = 1, C(u) = 0, u < t_0, A_i, B_i)}{q_1 P(E(t_0) = 1, E(t_1) = 0, C(t_0) = 1, C(u) = 0, u < t_0, A_i, B_i)}\right)^{-1}
\]
\[
\left(1 + \frac{P(C(t_0) = 1 \mid E(t_0) = 0, E(t_1) = 1, C(u) = 0, u < t_0, A_i, B_i)}{P(C(t_0) = 1 \mid E(t_0) = 1, E(t_1) = 0, C(u) = 0, u < t_0, A_i, B_i)} \right) \times \frac{P(E(t_0) = 0, E(t_1) = 1, C(u) = 0, u < t_0, A_i, B_i)}{P(E(t_0) = 1, E(t_1) = 0, C(u) = 0, u < t_0, A_i, B_i)}^{-1}
\]

\[
= \left(1 + \frac{\text{expit}(\beta_0 + B_i)}{\text{expit}(\beta_0 + B_i + \beta)} \frac{P(E(t_0) = 0 \mid E(t_1) = 1, C(u) = 0, u < t_0, A_i, B_i)}{P(E(t_0) = 1 \mid E(t_1) = 0, C(u) = 0, u < t_0, A_i, B_i)} \right)^{-1}
\]

\[
= \left(1 + \frac{\text{expit}(\beta_0 + B_i)}{\text{expit}(\beta_0 + B_i + \beta)} \frac{1 - \text{expit}(\alpha_0 + A_i + f(t_0))}{\text{expit}(\alpha_0 + A_i + f(t_0))} \frac{P(E(t_1) = 1, C(u) = 0, u < t_0, A_i, B_i)}{P(E(t_1) = 0, C(u) = 0, u < t_0, A_i, B_i)} \right)^{-1}
\]

\[
= \left(1 + \frac{\text{expit}(\beta_0 + B_i)}{\text{expit}(\beta_0 + B_i + \beta)} \frac{1 - \text{expit}(\alpha_0 + A_i + f(t_0))}{\text{expit}(\alpha_0 + A_i + f(t_0))} \frac{1 - \text{expit}(\beta_0 + B_i + \beta)}{1 - \text{expit}(\beta_0 + B_i)} \frac{\text{expit}(\alpha_0 + A_i + f(t_1))}{\text{expit}(\alpha_0 + A_i + f(t_1))} \right)^{-1}
\]

\[
= \left(1 + \frac{1 - \text{expit}(\beta_0 + B_i + \beta)}{\text{expit}(\beta_0 + B_i + \beta)} \frac{\text{expit}(\beta_0 + B_i)}{\text{expit}(\beta_0 + B_i) \ e^{\alpha_0 + A_i + f(t_1)}} \right)^{-1}
\]

\[
= \left(1 + \frac{e^{\beta_0 + B_i}}{e^{\beta_0 + B_i + \beta}} \ e^{f(t_1)-f(t_0)} \right)^{-1}
\]

\[
= \frac{e^{\beta + f(t_0)}}{e^{\beta + f(t_0)} + e^{f(t_1)}}
\]
Similarly one can show that:

\[ P(E(t_0) = 0, E(t_1) = 1 \mid \sum_{j=0}^{1} E(t_j) = 1, C(t_0) = 1, C(u) = 0, u < t_0, A_i, B_i, S = 1) = \frac{e^{f(t_1)}}{e^{f(t_1)} + e^{\beta + f(t_0)}} \]

\[ \blacksquare \]

### 3.4.2 Controls:

**lemma 3.4.2.**

(1) \( P(E(t_0) = 1, E(t_1) = 0 \mid \sum_{j=0}^{1} E(t_j) = 1, C(t_0) = 0, C(u) = 0, u < t_0, A_i, B_i, S = 1) = \frac{e^{f(t_0)}}{e^{f(t_0)} + e^{f(t_1)}} \)

(2) \( P(E(t_0) = 0, E(t_1) = 1 \mid \sum_{j=0}^{1} E(t_j) = 1, C(t_0) = 0, C(u) = 0, u < t_0, A_i, B_i, S = 1) = \frac{e^{f(t_1)}}{e^{f(t_0)} + e^{f(t_1)}} \)

**Proof.** Completely analogous to the proof for the cases except that the \( \beta \)-terms equals out for the controls. \( \blacksquare \)

### 3.5 Likelihood-calculations for 2 reference periods

#### 3.5.1 Cases:

**lemma 3.5.1.**

(1) \( P(E(t_0) = 1, E(t_1) = 0, E(t_2) = 0 \mid \sum_{j=0}^{2} E(t_j) = 1, C(t_0) = 1, C(u) = 0, u < t_0, A_i, B_i, S = 1) \)

\[ = \frac{e^{f(t_0)}}{e^{\beta + f(t_0)} + e^{f(t_1)} + e^{f(t_2)}} \]

(2) \( P(E(t_0) = 0, E(t_1) = 1, E(t_2) = 0 \mid \sum_{j=0}^{2} E(t_j) = 1, C(t_0) = 1, C(u) = 0, u < t_0, A_i, B_i, S = 1) \)

\[ = \frac{e^{f(t_1)}}{e^{\beta + f(t_0)} + e^{f(t_1)} + e^{f(t_2)}} \]

(3) \( P(E(t_0) = 0, E(t_1) = 0, E(t_2) = 1 \mid \sum_{j=0}^{2} E(t_j) = 1, C(t_0) = 1, C(u) = 0, u < t_0, A_i, B_i, S = 1) \)

\[ = \frac{e^{f(t_2)}}{e^{\beta + f(t_0)} + e^{f(t_1)} + e^{f(t_2)}} \]
Proof. The proofs for (1)-(6) are similar. We outline (5) in details.

\[
P(E(t_0) = 1, E(t_1) = 1, E(t_2) = 0 \bigg| \sum_{j=0}^{2} E(t_j) = 2, C(t_0) = 1, C(u) = 0, u < t_0, A_i, B_i, S = 1) = \\
\frac{e^{\beta+f(t_0)+f(t_1)}}{e^{\beta+f(t_0)+f(t_1)} + e^{\beta+f(t_0)+f(t_2)} + e^{f(t_1)+f(t_2)}}
\]

\[
P(E(t_0) = 1, E(t_1) = 0, E(t_2) = 1 \bigg| \sum_{j=0}^{2} E(t_j) = 2, C(t_0) = 1, C(u) = 0, u < t_0, A_i, B_i, S = 1) = \\
\frac{e^{\beta+f(t_0)+f(t_2)}}{e^{\beta+f(t_0)+f(t_1)} + e^{\beta+f(t_0)+f(t_2)} + e^{f(t_1)+f(t_2)}}
\]

\[
P(E(t_0) = 0, E(t_1) = 1, E(t_2) = 1 \bigg| \sum_{j=0}^{2} E(t_j) = 2, C(t_0) = 1, C(u) = 0, u < t_0, A_i, B_i, S = 1) = \\
\frac{e^{f(t_1)+f(t_2)}}{e^{\beta+f(t_0)+f(t_1)} + e^{\beta+f(t_0)+f(t_2)} + e^{f(t_1)+f(t_2)}}
\]

Proof. The proofs for (1)-(6) are similar. We outline (5) in details.

\[
P(E(t_0) = 1, E(t_1) = 0, E(t_2) = 1 \bigg| \sum_{j=0}^{1} E(t_j) = 2, C(t_0) = 1, C(u) = 0, u < t_0, A_i, B_i, S = 1) = \\
\frac{1 + \frac{P(E(t_0) = 1, E(t_1) = 1, E(t_2) = 0, C(t_0) = 1, C(u) = 0, u < t_0, A_i, B_i, S = 1)}{P(E(t_0) = 1, E(t_1) = 1, E(t_2) = 0, C(t_0) = 1, C(u) = 0, u < t_0, A_i, B_i, S = 1)}}{P(E(t_0) = 1, E(t_1) = 1, E(t_2) = 1, C(t_0) = 1, C(u) = 0, u < t_0, A_i, B_i, S = 1)} - 1
\]

\[
= \left(1 + \frac{P(C(t_0) = 1 \bigg| E(t_0) = 1, E(t_1) = 1, E(t_2) = 0, C(u) = 0, u < t_0, A_i, B_i)}{P(C(t_0) = 1 \bigg| E(t_0) = 1, E(t_1) = 0, E(t_2) = 1, C(u) = 0, u < t_0, A_i, B_i)} \right) \times \frac{P(E(t_0) = 1, E(t_1) = 1, E(t_2) = 1, C(u) = 0, u < t_0, A_i, B_i)}{P(E(t_0) = 1, E(t_1) = 1, E(t_2) = 0, C(u) = 0, u < t_0, A_i, B_i)}
\]

\[
+ \frac{P(C(t_0) = 1 \bigg| E(t_0) = 0, E(t_1) = 1, E(t_2) = 1, C(u) = 0, u < t_0, A_i, B_i)}{P(C(t_0) = 1 \bigg| E(t_0) = 1, E(t_1) = 0, E(t_2) = 1, C(u) = 0, u < t_0, A_i, B_i)}
\]

\[
+ \frac{P(C(t_0) = 1 \bigg| E(t_0) = 0, E(t_1) = 1, E(t_2) = 1, C(u) = 0, u < t_0, A_i, B_i)}{P(C(t_0) = 1 \bigg| E(t_0) = 1, E(t_1) = 0, E(t_2) = 1, C(u) = 0, u < t_0, A_i, B_i)}
\]
$$\begin{align*}
&\times \frac{P(E(t_0) = 0, E(t_1) = 1, E(t_2) = 1, C(u) = 0, u < t_0, A_i, B_i)}{P(E(t_0) = 1, E(t_1) = 0, E(t_2) = 1, C(u) = 0, u < t_0, A_i, B_i)}^{-1} \\
&= \left(1 + \frac{\expit(\beta_0 + B_i + \beta)}{\expit(\beta_0 + B_i + \beta)} \frac{P(E(t_0) = 1, E(t_1) = 1, E(t_2) = 0, C(u) = 0, u < t_0, A_i, B_i)}{P(E(t_0) = 1, E(t_1) = 0, E(t_2) = 1, C(u) = 0, u < t_0, A_i, B_i)} \right. \\
&+ \left. \frac{\expit(\beta_0 + B_i + \beta)}{\expit(\beta_0 + B_i + \beta)} \frac{P(E(t_0) = 0, E(t_1) = 1, E(t_2) = 1, C(u) = 0, u < t_0, A_i, B_i)}{P(E(t_0) = 1, E(t_1) = 0, E(t_2) = 1, C(u) = 0, u < t_0, A_i, B_i)} \right)^{-1} \\
&= \left(1 + \frac{P(E(t_0) = 1 \mid E(t_1) = 1, E(t_2) = 0, C(u) = 0, u < t_0, A_i, B_i)}{P(E(t_0) = 1 \mid E(t_1) = 0, E(t_2) = 1, C(u) = 0, u < t_0, A_i, B_i)} \right. \\
&\times \left. \frac{P(E(t_1) = 1, E(t_2) = 0, C(u) = 0, u < t_0, A_i, B_i)}{P(E(t_1) = 0, E(t_2) = 1, C(u) = 0, u < t_0, A_i, B_i)} \right)^{-1} \\
&= \left(1 + \frac{P(E(t_1) = 1, E(t_2) = 0, C(u) = 0, u < t_0, A_i, B_i)}{P(E(t_1) = 0, E(t_2) = 1, C(u) = 0, u < t_0, A_i, B_i)} + \frac{\expit(\beta_0 + B_i)}{\expit(\beta_0 + B_i + \beta)} \\
&\times \frac{1 - \expit(\alpha_0 + A_i + f(t_0))}{\expit(\alpha_0 + A_i + f(t_0))} \frac{P(E(t_1) = 1, E(t_2) = 1, C(u) = 0, u < t_0, A_i, B_i)}{P(E(t_1) = 0, E(t_2) = 1, C(u) = 0, u < t_0, A_i, B_i)} \right)^{-1} \\
&= \left(1 + \frac{P(C(t_1) = 0 \mid E(t_1) = 1, E(t_2) = 0, C(u) = 0, u < t_1, A_i, B_i)}{P(C(t_1) = 0 \mid E(t_1) = 0, E(t_2) = 1, C(u) = 0, u < t_1, A_i, B_i)} \right. \\
&\times \left. \frac{P(E(t_1) = 1, E(t_2) = 0, C(u) = 0, u < t_1, A_i, B_i)}{P(E(t_1) = 0, E(t_2) = 1, C(u) = 0, u < t_1, A_i, B_i)} + \frac{\expit(\beta_0 + B_i)}{\expit(\beta_0 + B_i + \beta)} \\
&\times \frac{1 - \expit(\alpha_0 + A_i + f(t_0))}{\expit(\alpha_0 + A_i + f(t_0))} \frac{P(C(t_1) = 0 \mid E(t_1) = 1, E(t_2) = 1, C(u) = 0, u < t_1, A_i, B_i)}{P(C(t_1) = 0 \mid E(t_1) = 0, E(t_2) = 1, C(u) = 0, u < t_1, A_i, B_i)} \right)^{-1} \\
&\times \left. \frac{P(E(t_1) = 1, E(t_2) = 1, C(u) = 0, u < t_1, A_i, B_i)}{P(E(t_1) = 0, E(t_2) = 1, C(u) = 0, u < t_1, A_i, B_i)} \right)^{-1}
\end{align*}
\[
= \left( 1 + \frac{1 - \expit(\beta_0 + B_i + \beta)}{1 - \expit(\beta_0 + B_i)} \cdot \frac{P\left( E(t_1) = 1, E(t_2) = 0, C(u) = 0, u < t_1, A_i, B_i \right)}{P\left( E(t_1) = 0, E(t_2) = 1, C(u) = 0, u < t_1, A_i, B_i \right)} \right.
\]
\[
+ \frac{\expit(\beta_0 + B_i)}{1 - \expit(\beta_0 + B_i + \beta)} \cdot \frac{1 - \expit\left(\beta_0 + B_i + \beta\right)}{1 - \expit(\beta_0 + B_i)} \cdot \frac{1 - \expit(\beta_0 + B_i + \beta)}{1 - \expit(\beta_0 + B_i) \expit\left(\alpha_0 + A_i + f(t_0)\right)} \cdot \frac{1 - \expit(\beta_0 + B_i + \beta)}{1 - \expit(\beta_0 + B_i) \expit\left(\alpha_0 + A_i + f(t_0)\right)}
\]
\[
\left. \times \frac{P\left( E(t_1) = 1, E(t_2) = 1, C(u) = 0, u < t_1, A_i, B_i \right)}{P\left( E(t_1) = 0, E(t_2) = 1, C(u) = 0, u < t_1, A_i, B_i \right)} \right)^{-1}
\]
\[\begin{align*}
&= \left(1 + e^{\alpha_0 + A_i + f(t_1)} \frac{1 - \expit(\alpha_0 + A_i + f(t_2))}{\expit(\alpha_0 + A_i + f(t_2))} + e^{-\beta} e^{\alpha_0 + A_i + f(t_1)} \right)^{-1} \\
&= \left(1 + e^{f(t_1) - f(t_2)} + e^{-\beta + f(t_1) - f(t_0)} \right)^{-1} \\
&= \frac{e^{\beta + f(t_0) + f(t_2)}}{e^{\beta + f(t_0) + f(t_2)} + e^{\beta + f(t_0) + f(t_1) + f(t_2)}}
\end{align*}\]

\[\square\]

3.5.2 Controls:

**Lemma 3.5.2.**

1. \[P\{ E(t_0) = 1, E(t_1) = 0, E(t_2) = 0 \mid \sum_{j=0}^{2} E(t_j) = 1, C(t_0) = 0, C(u) = 0, \ u < t_0, A_i, B_i, S = 1 \} = \frac{e^{f(t_0)}}{e^{f(t_0)} + e^{f(t_1)} + e^{f(t_2)}}\]

2. \[P\{ E(t_0) = 0, E(t_1) = 1, E(t_2) = 0 \mid \sum_{j=0}^{2} E(t_j) = 1, C(t_0) = 0, C(u) = 0, \ u < t_0, A_i, B_i, S = 1 \} = \frac{e^{f(t_1)}}{e^{f(t_0)} + e^{f(t_1)} + e^{f(t_2)}}\]

3. \[P\{ E(t_0) = 0, E(t_1) = 0, E(t_2) = 1 \mid \sum_{j=0}^{2} E(t_j) = 1, C(t_0) = 0, C(u) = 0, \ u < t_0, A_i, B_i, S = 1 \} = \frac{e^{f(t_2)}}{e^{f(t_0)} + e^{f(t_1)} + e^{f(t_2)}}\]

4. \[P\{ E(t_0) = 1, E(t_1) = 1, E(t_2) = 0 \mid \sum_{j=0}^{2} E(t_j) = 2, C(t_0) = 0, C(u) = 0, \ u < t_0, A_i, B_i, S = 1 \} = \frac{e^{f(t_0) + f(t_1)}}{e^{f(t_0) + f(t_1)} + e^{f(t_0) + f(t_2)} + e^{f(t_1) + f(t_2)}}\]

5. \[P\{ E(t_0) = 1, E(t_1) = 0, E(t_2) = 1 \mid \sum_{j=0}^{2} E(t_j) = 2, C(t_0) = 0, C(u) = 0, \ u < t_0, A_i, B_i, S = 1 \} = \frac{e^{f(t_0) + f(t_2)}}{e^{f(t_0) + f(t_1)} + e^{f(t_0) + f(t_2)} + e^{f(t_1) + f(t_2)}}\]
(6) \[ P \left( E(t_0) = 0, E(t_1) = 1, E(t_2) = 1 \mid \sum_{j=0}^{2} E(t_j) = 2, C(t_0) = 0, C(u) = 0, u < t_0, A_i, B_i, S = 1 \right) \]

\[ = \frac{e^{f(t_1) + f(t_2)}}{e^{f(t_0) + f(t_1)} + e^{f(t_0) + f(t_2)} + e^{f(t_1) + f(t_2)}} \]

**Proof.** As with one reference period the proofs for the cases and the controls are similar, again all \( \beta \)-terms disappears in the calculations for the controls.

\[ \square \]
Research Reports available from Department of Biostatistics

http://www.pubhealth.ku.dk/bs/publikationer

10/1  Andersen, P.K. & Skrondal, A. “Biological” interaction from a statistical point of view.


10/3  Parner, E.T. & Andersen, P.K. Regression analysis of censored data using pseudo-observations.

10/4  Nielsen, T. & Kreiner, S. Course Evaluation and Development: What can Learning Styles Contribute?

10/5  Lange, T. & Hansen, J.V. Direct and Indirect Effects in a Survival Context.

10/6  Andersen, P.K. & Keiding, N. Interpretability and importance of functionals in competing risks and multi-state models.

10/7  Gerds, T.A., Kattan, M.W., Schumacher, M. & Yu, C. Estimating a time-dependent concordance index for survival prediction models with covariate dependent censoring.


11/1  Kreiner, S. Is the foundation under PISA solid? A critical look at the scaling model underlying international comparisons of student attainment.

11/2  Andersen, P.K. Competing risks in epidemiology: Possibilities and pitfalls.

11/3  Holst, K.K. Model diagnostics based on cumulative residuals: The R-package gof.

11/5 Holst, K.K., Budtz-Jørgensen, E. & Knudsen, G.M. A latent variable model with mixed binary and continuous response variables.


12/01 Keiding, N., Hansen, O.K.H., Sørensen, D.N. & Slama, R. The current duration approach to estimating time to pregnancy.

12/02 Andersen, P.K. A note on the decomposition of number of life years lost according to causes of death.

12/03 Martinussen, T. & Vansteelandt, S. A note on collapsibility and confounding bias in Cox and Aalen regression models.


12/05 Martinussen, T. & Pipper, C.B. Estimation of Causal Odds of Concordance based on the Aalen additive model.

12/06 Binder, N., Gerds, T.A. & Andersen, P.K. Pseudo-observations for competing risks with covariate dependent censoring.

12/07 Keiding, N. & Clayton, D. Standardization and control for confounding in observational studies: a historical perspective.

12/08 Hilden, J. & Gerds, T.A. Evaluating the impact of novel biomarkers: Do not rely on IDI and NRI.

