Supplementary material to 'Cause-specific survival analysis with misclassified cause-of-failure: Estimation with Cox Models'

1 Additional topics from the small illustration

1.1 Simulation setting

We consider simulated data for a clinical trial accruing 2543 patients over 2 years time, with an additional 4 years of follow-up. The patients are randomized to treatment ($X = 1$) or control ($X = 0$) with equal probability. Two causes of death occur with constant cause-specific baseline hazard for cause 1, h_1, equal to 0.04 and for cause 0, h_0, equal to 0.06.

The main interest lies in the treatment effect on the cause-1-specific hazard. We consider settings where this log cause-1-specific hazard ratio is -0.3 (treatment lowers the hazard if interest) or 0 (no effect), and where the log cause-0-specific hazard ratio is 0.1 (competing risk hazard increases under treatment) or 0 (no effect). The setting without any effects is called the global null setting, when both effect are present this is called the full alternative setting. When an effect is present on one event type, but not the other, this is called the partial alternative setting. The full alternative setting was already described in the main text.

We consider two types of analysis: an all-cause analysis on the one hand, and the two cause-specific analyses on the other. We always use Cox models to obtain effect estimates and Wald tests.

1.2 Sample size calculation

We consider two possible treatments $X \in \{0, 1\}$ with equal randomization probabilities, in a trial with 2 years staggered accrual and an additional 4 years of follow-up. Two causes of death occur with constant cause-specific baseline hazard for cause 1 $h_1 = 0.04$ and for cause 0 $h_0 = 0.06$. Treatment affects both the cause-1-specific hazard, lowering it by a factor $\exp(\phi X)$=$\exp(-0.3X)$, and the cause-0-specific hazard, increasing it with a factor $\exp(\rho X)$=$\exp(0.1X)$. We size the study to yield a power of 80% for detecting the effect on the cause-1-specific hazard at the 5% significance level.

To compute the needed sample size we use the approach from Schoenfeld,[1] which here becomes:

$$n = \frac{4(z_{1-\frac{\alpha}{2}} + z_{1-\beta})^2}{P \log(\exp(\phi))^2}$$

In this formula, the z's indicate quantiles from the standard normal distribution, and P is the overall probability for any individual in the study of seeing an event of interest (type 1). This

misclassification	all-cause	cause-specific cause 1	cause 0
no	0.0497	0.0527	0.0506
yes	0.0497	0.0522	0.0479

Table 1: Type I error rates under the global null ($\phi = 0$, $\rho = 0$). Misclassification rates are $p_1 = 0.6$ and $p_0 = 0.1$ (standard error on estimates below 0.005).

overall probability is computed as the mean probability in the both treatment groups:

$$P = \frac{P_{X=0} + P_{X=1}}{2},$$

where

$$P_{X=0} = \frac{h_1}{h_1 + h_0}\left(1 - \frac{e^{-(h_1+h_0)f} + e^{-(h_1+h_0)(a+f)}}{(h_1 + h_0)a}\right)$$

and

$$P_{X=1} = \frac{h_1 e^{\phi}}{h_1 e^{\phi} + h_0 e^{\rho}}\left(1 - \frac{e^{-(h_1 e^{\phi}+h_0 e^{\rho})f} + e^{-(h_1 e^{\phi}+h_0 e^{\rho})(a+f)}}{(h_1 e^{\phi} + h_0 e^{\rho})a}\right)$$

These expressions are appropriate for trials with uniform staggered accrual where only administrative censoring occurs, when the hazards are constant (see e.g. Pintilie[2]).

1.3 Type I error

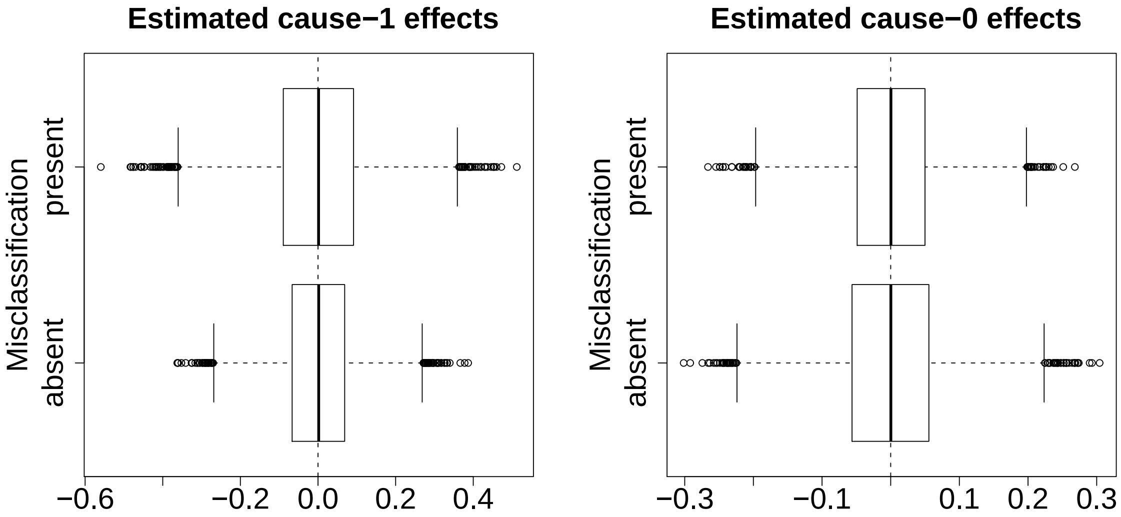
1.3.1 Type I error under the global null

First, we provide details on the behaviour under the global null ($\phi = \rho = 0$). Generally speaking, the effect of misclassification is one of mixing of cause-specific hazards within treatment groups. Since there is no difference in either type of mortality for treated versus untreated, the misclassified mortalities will also be indifferent if misclassification occurs equally between the treated and untreated group. Because of this, we only consider one set of misclassification probabilities, namely $p_1 = 0.6$ and $p_0 = 0.1$. Table 1 shows the results from 10,000 simulations for this setting. In each simulation step, there are around 350 true type-1 events and 640 type-0 events, which are misclassified to yield around 204 type-1 observations and 786 type-0 observations. All tests respect the 5% significance level.

The unbiasedness of the estimators is illustrated in Figure 1, which shows the distribution of simulation estimates. The true value of 0 is indicated by the vertical striped line. The difference in variance is due to the net decrease in type-1 observations under misclassification, and the net increase in type-0 observations.

1.3.2 Type I error under treatment-specific misclassification rates

In some settings, it is possible that the misclassification rates are different between treatment groups, e.g. when one treatment suppresses certain symptoms, identification of the cause-of-death may be more difficult. In such situations, the type I error of a cause-specific analysis can be increased. We redo the previous simulation study, now assuming that $p_1 = 0.6$ and $p_0 = 0.1$ in the control group, but $p_1 = 0.4$ and $p_0 = 0.1$ in the treatment group.



(a) Cause-1-specific effect

(b) Cause-0-specific effect

Figure 1: *Distribution of effect estimates under the global null in the absence (bottom) and presence (top) of misclassification. The true effect is indicated by a vertical striped line.*

misclassification	all-cause	cause-specific	
		cause 1	cause 0
no	0.0497	0.0533	0.0525
yes	0.0497	0.7036	0.3130

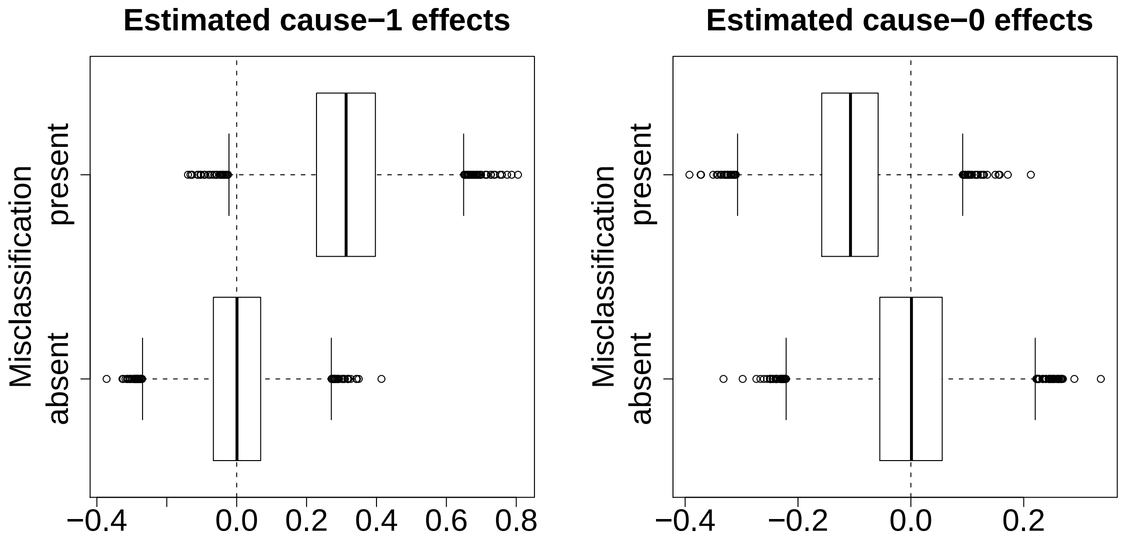
Table 2: Type I error rates under the global null ($\phi = 0$, $\rho = 0$), with differential misclassification rates between treatment groups (standard error on estimates below 0.005).

Table 2 illustrates the inflation of the type I error rate that results from such differential misclassification probabilities. This is due to a treatment-specific mixing of cause-specific hazards. Figure 2 shows the bias that is induced in the cause-specific analyses in this case. This is limited in our chosen settings, as the difference between $p_1 = 0.4$ and $p_1 = 0.6$ is rather small and there are fewer cause-1 events than cause-0 events.

1.3.3 Type I error under the partial alternative

It is very well possible that the true setting is a partial alternative setting, where a treatment affects the hazard of one type of event, but not the other type. We simulate the setting where the hazard of the competing risk depends on the treatment ($h_0 = 0.06 \exp(0.1X)$), but not the hazard of the event of interest ($h_1 = 0.04$). To clearly show the effect on the type I error for the event of interest, the misclassification probability p_0 is now 0.3, while p_1 is still 0.6. Table 3 shows the results from 10,000 simulations for this setting.

This shows how the type I error for the effect on the event of interest is inflated, while the power for the effect on the competing risk becomes smaller. This is the result of the carry-over of the competing risk effect onto the event of interest. This bias induced in the estimators is illustrated in Figure 3. While the bias is limited, one must realize that the two effects $\phi = 0$ and $\rho = 0.1$ are fairly similar, and the bias will increase as these two diverge.



(a) Cause-1-specific effect

(b) Cause-0-specific effect

Figure 2: *Distribution of effect estimates under the global null in the absence (bottom) and presence (top) of misclassification. The misclassification rate for type-1 events now depends on the treatment group. The true effect is indicated by a vertical striped line.*

	all-cause	cause-specific	
misclassification		cause 1	cause 0
no	0.1693	0.0521	0.2481
yes	0.1693	0.0810	0.1361

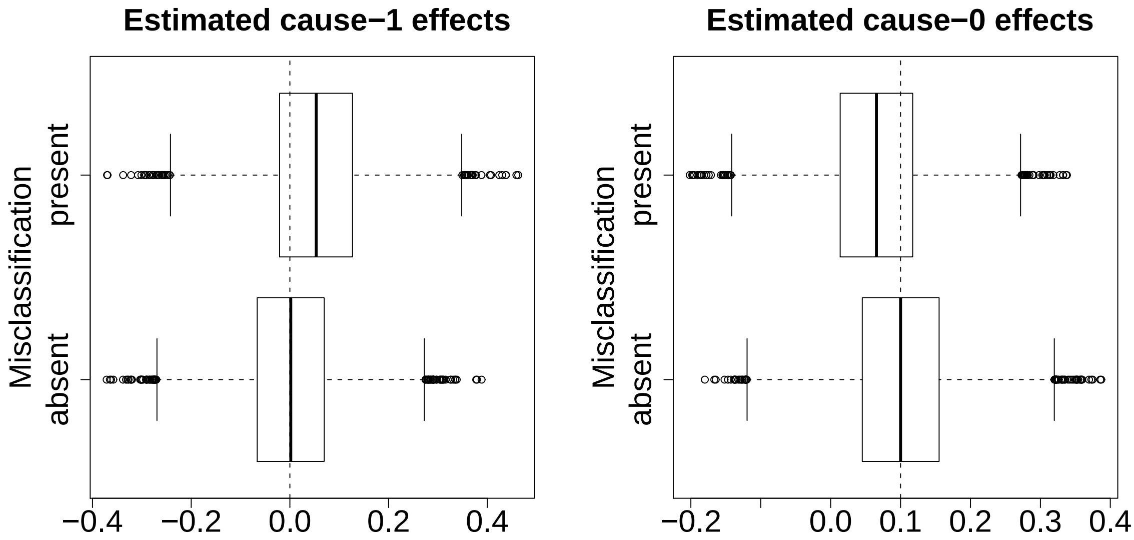
Table 3: Type I error rates under the partial alternative ($\phi = 0$, $\rho = 0.1$). Misclassification rates are $p_1 = 0.6$ and $p_0 = 0.3$ (standard error on estimates below 0.005).

1.4 Summary of the simulation results

The all-cause analysis always respects the type I error rate, since the diagnosis itself is not considered in the analysis. The cause-specific analyses however, respect the assumed type I error rate only under the global null when misclassification is the same in both treatment groups. Every deviation from this leads to bias and to an increased type I error: under a full alternative scenario, but also when a treatment effect is present on only one event type, and even under the global null with differential misclassification errors between groups. The rationale behind all this is that within each treatment group, the true cause-specific hazards are mixed to form observed hazards, and subsequent comparison of treatment groups leads to false conclusions.

2 Comprehensive study of the proposed analysis in terms of MSE

The previous evaluation of our methodology in terms of type I error and power gives only a limited view of what is truly going on. To better understand the underlying dynamics of these changes in rejection rates, a comprehensive simulation study of the MSE was performed.



(a) Cause-1-specific effect

(b) Cause-0-specific effect

Figure 3: *Distribution of effect estimates under the partial alternative in the absence (bottom) and presence (top) of misclassification. The true effect is indicated by a vertical striped line.*

We simulate a clinical trial with 2 years of uniform accrual followed by 4 additional years of follow-up. Censoring was administrative 6 years after start of the study. Individuals are randomized over a binary treatment with probability 0.5. Two causes of death exist, but we hold the overall survival hazard constant at 0.6 in the control group. The cause-specific hazards are constant and in the control group equal to $0.6/(1+\exp(\xi))$ (cause 0) and $0.6/(1+\exp(-\xi))$ (cause 1). For the treatment group, the cause-0-specific hazard is $0.6\exp(\rho)/(1+\exp(\xi))$, the cause-1-specific hazard is $0.6\exp(\phi)/(1+\exp(-\xi))$. The cause of death is misdiagnosed with misclassification probabilities p_0 and p_1.

In the simulation study, the parameter set (ϕ, ρ, ξ, p_0, p_1) takes on all possible combinations with:

- $\phi \in \{0, 1\}$
- $\rho \in \{-0.5, 0, 0.5\}$
- $\xi \in \{-1, 0, 1\}$
- $p_0 \in \{0.01, 0.2, 0.5\}$
- $p_1 \in \{0.01, 0.2, 0.5\}$

for a total of 162 parameter settings. Typically, the censoring rate was roughly 10%. Each of these settings was simulated using sample sizes of 200, 500, 1000, 5000 and 10000 (as lower sample sizes can lead to unstable results in individual simulations). At each sample size of each setting, 1000 simulations were performed. In each simulation, we apply a standard cause-1-specific Cox model, and we apply our method assuming a constant ξ.

2.1 A setting where our method is useful

The general idea is that the new method sacrifices a certain loss in terms of variance to avoid the bias that is introduced in standard cause-specific analyses in the presence of misclassification. With increasing sample size however, this variance gets smaller and the bias dominates the MSE of the analyses, making our method increasingly appealing.

Figure 4 illustrates this for the setting with $\phi = 1$, $\rho = -0.5$ and $\xi = -1$. Here, the opposing direction of the cause-specific effects leads to considerable bias in the naive, unadjusted analysis. This bias is increased by the choice for ξ, which ensures that the (more prevalent) competing risk events pollute the pool of events of interest more severely.

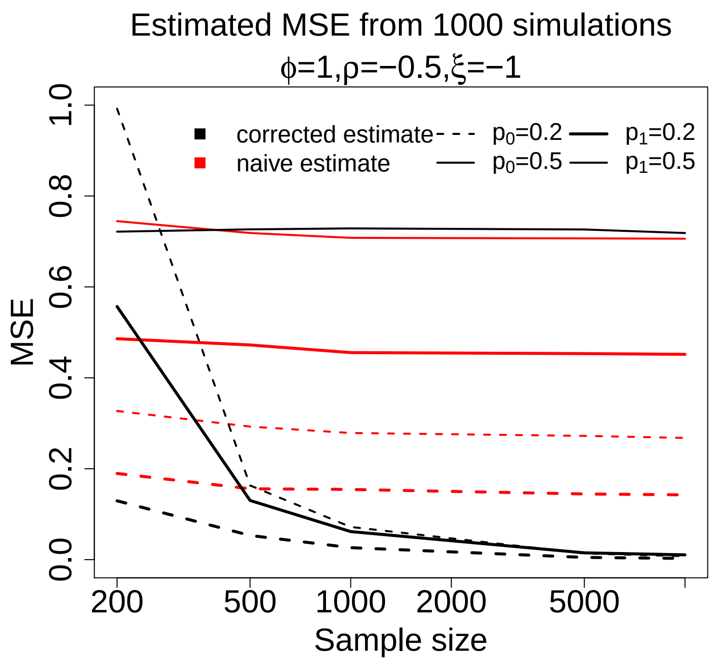


Figure 4: *Illustration of the increasing advantage in terms of MSE with increasing sample size. The naive analysis is represented by red lines, the adjusted by black ones. The level of p_0 is represented by the line style, p_1 is represented by line thickness.*

The analyses have higher MSE at lower sample size, but as n increases the variance component reduces to 0, while the bias component in the naive analysis remains constant. At a moderate sample size of 500 the adjusted analysis already performs much better than the naive one.

One exception is the analysis with $p_0 = p_1 = 0.5$. Here the diagnosis is completely at random, it contains no information whatsoever. Both analyses are then similar to an all-cause analysis, but the naive, standard analysis doesn't use all of the available data (it uses only those diagnosed as being of interest). Our method on the other hand becomes similar to a weighted all-cause analysis, and takes all events into account. This explains the somewhat higher initial variance of the standard cause-specific analysis. It is however clear that if information on the cause of death is completely absent, our method also fails to remove the bias. Intuitively, it should be clear that performing any type of cause-specific analysis in this setting is destined for failure.

Still, to clarify the gain in variance, we look at some more simulations without information on the cause of death ($p_0 = p_1 = 0.5$). From the simulations, the decreased variance (our method compared to the standard cause-specific analysis) that results from this is clear (Figure 5).

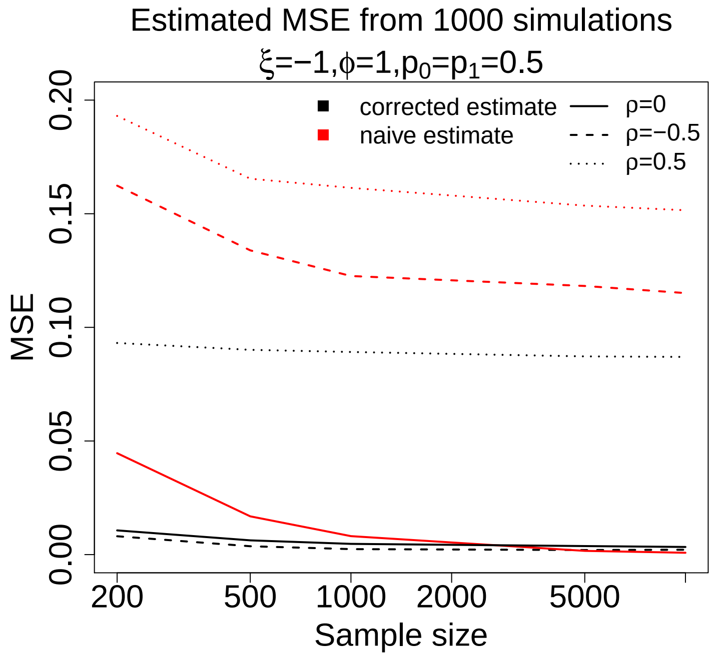


Figure 5: *Illustration of the behaviour when* $p_0 = p_1 = 0.5$*. On the left hand side, one sees how the standard cause-specific analysis (in red) has an increased variance. The line style now indicates* ρ*,* ξ *is fixed at -1 and* ϕ *at 0.*

2.2 Two settings without gain

In some settings, our method has less or no added value compared to the standard analysis, or the impact of using our method is at least ambiguous. In the presence of misclassification, this is the case in contrived examples where the treatment-specific cause-specific hazards are exactly balanced in such a way that no bias occurs in a naive analysis.

However, it also occurs when the information on the cause of death is quasi perfect ($p_0 = p_1 \sim 0$). The partial loglikelihood l then separates into two parts, one equal to the partial loglikelihood used in a standard cause 1-specific setting, and one equal to the partial loglikelihood used in a standard cause 0-specific setting. This means that our method then makes optimal use of the information on the cause-of-death, and becomes exactly the same as a standard cause-specific analysis.

When the information on the cause-of-death is quasi-perfect (e.g. $p_0 = p_1 = 0.01$), the bias in the standard cause-specific analysis will be minimal, while its variance is optimal. While not the same, Figure 6 illustrates that the variance of our method is only slightly increased (due to the explicit estimation of the nuisance parameter ξ).

Finally, under the global null ($\phi = \rho = 0$), no bias occurs in the standard cause-specific analysis. Because of its complexity, the adjusted analysis again shows increased variance (Figure 7). The method therefore seems less appealing. However, we do note that the variance of our method is estimated correctly, meaning that the type I error rate is not inflated and tests are unbiased.

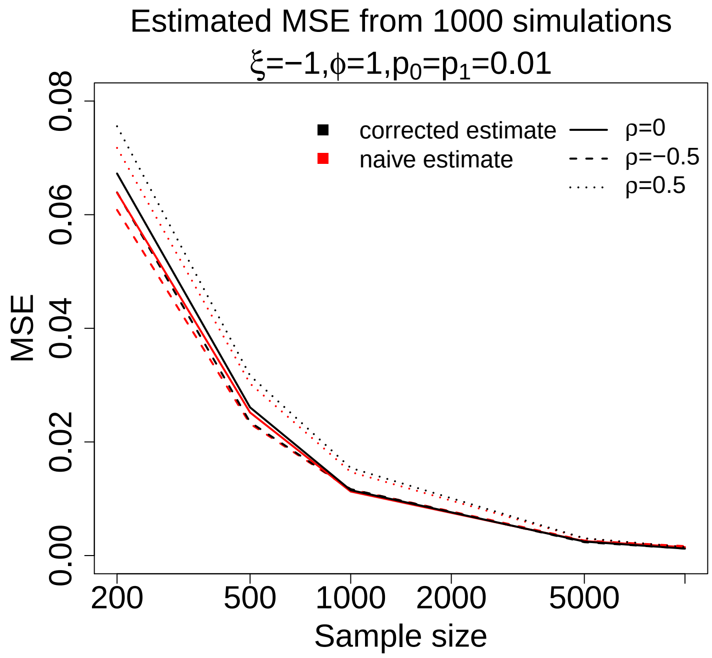


Figure 6: *Illustration of the behaviour when* $p_0 = p_1 = 0.01$. *On the left hand side, one sees how our method (black lines) has an only slightly increased variance. The line style indicates* ρ, ξ *is fixed at -1 and* ϕ *at 0.*

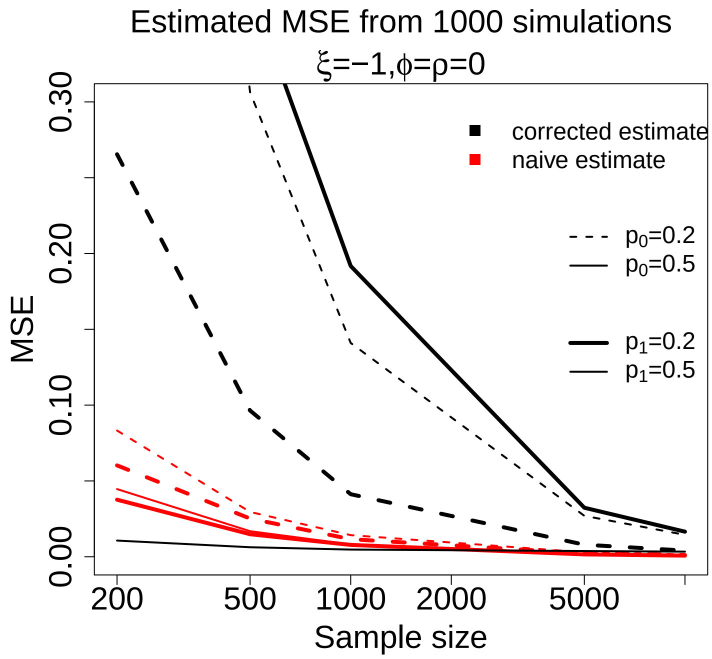


Figure 7: *Illustration of the behaviour under the global null* $\phi = \rho = 0$. *While both methods are unbiased, our method (black lines) has increased variance. The line style and thickness indicate* p_0 *and* p_1, *respectively.*

2.3 Listing of simulation results

To allow the reader to make his own comparisons, if desired, we tabulate the MSE values from the simulations for all simulation settings, first for a naive, standard cause-specific analysis (Table 4), next for our adjusted analysis (Table 5).

Table 4: *Empirical MSE's for an unadjusted cause-specific analysis for the 162 parameter settings, based on 1000 simulations.*

	parameter settings					empirical MSE				
setting	ϕ	ρ	ξ	p_0	p_1	n=200	n=500	n=1,000	n=5,000	n=10,000
1	0	-0.50	-1.00	0.01	0.01	0.0777	0.0260	0.0133	0.0031	0.0015
2	0	-0.50	-1.00	0.01	0.20	0.0873	0.0353	0.0170	0.0034	0.0018
3	0	-0.50	-1.00	0.01	0.50	0.1501	0.0533	0.0292	0.0057	0.0033
4	0	-0.50	-1.00	0.20	0.01	0.0705	0.0439	0.0317	0.0239	0.0236
5	0	-0.50	-1.00	0.20	0.20	0.0896	0.0574	0.0424	0.0331	0.0318
6	0	-0.50	-1.00	0.20	0.50	0.1331	0.0846	0.0676	0.0565	0.0540
7	0	-0.50	-1.00	0.50	0.01	0.1103	0.0803	0.0747	0.0678	0.0678
8	0	-0.50	-1.00	0.50	0.20	0.1192	0.1016	0.0880	0.0828	0.0814
9	0	-0.50	-1.00	0.50	0.50	0.1624	0.1339	0.1226	0.1182	0.1151
10	0	-0.50	0.00	0.01	0.01	0.0379	0.0151	0.0077	0.0016	0.0008
11	0	-0.50	0.00	0.01	0.20	0.0465	0.0195	0.0092	0.0020	0.0010
12	0	-0.50	0.00	0.01	0.50	0.0831	0.0314	0.0163	0.0034	0.0017
13	0	-0.50	0.00	0.20	0.01	0.0379	0.0182	0.0116	0.0061	0.0054
14	0	-0.50	0.00	0.20	0.20	0.0490	0.0235	0.0140	0.0085	0.0076
15	0	-0.50	0.00	0.20	0.50	0.0715	0.0374	0.0267	0.0164	0.0154
16	0	-0.50	0.00	0.50	0.01	0.0448	0.0327	0.0264	0.0209	0.0206
17	0	-0.50	0.00	0.50	0.20	0.0605	0.0397	0.0334	0.0287	0.0283
18	0	-0.50	0.00	0.50	0.50	0.0938	0.0672	0.0562	0.0502	0.0490
19	0	-0.50	1.00	0.01	0.01	0.0306	0.0111	0.0053	0.0011	0.0006
20	0	-0.50	1.00	0.01	0.20	0.0371	0.0140	0.0065	0.0013	0.0007
21	0	-0.50	1.00	0.01	0.50	0.0578	0.0211	0.0107	0.0022	0.0010
22	0	-0.50	1.00	0.20	0.01	0.0272	0.0110	0.0060	0.0018	0.0013
23	0	-0.50	1.00	0.20	0.20	0.0321	0.0128	0.0082	0.0023	0.0018
24	0	-0.50	1.00	0.20	0.50	0.0513	0.0242	0.0119	0.0052	0.0039
25	0	-0.50	1.00	0.50	0.01	0.0285	0.0137	0.0086	0.0051	0.0045
26	0	-0.50	1.00	0.50	0.20	0.0353	0.0178	0.0116	0.0069	0.0063
27	0	-0.50	1.00	0.50	0.50	0.0531	0.0277	0.0210	0.0146	0.0136
28	0	0.00	-1.00	0.01	0.01	0.0768	0.0300	0.0155	0.0030	0.0016
29	0	0.00	-1.00	0.01	0.20	0.1051	0.0353	0.0192	0.0037	0.0019
30	0	0.00	-1.00	0.01	0.50	0.1669	0.0678	0.0292	0.0055	0.0031
31	0	0.00	-1.00	0.20	0.01	0.0523	0.0213	0.0096	0.0022	0.0011
32	0	0.00	-1.00	0.20	0.20	0.0602	0.0252	0.0116	0.0022	0.0011
33	0	0.00	-1.00	0.20	0.50	0.0832	0.0295	0.0143	0.0029	0.0016
34	0	0.00	-1.00	0.50	0.01	0.0311	0.0143	0.0066	0.0014	0.0006
35	0	0.00	-1.00	0.50	0.20	0.0376	0.0149	0.0078	0.0016	0.0007
36	0	0.00	-1.00	0.50	0.50	0.0446	0.0168	0.0081	0.0016	0.0008
37	0	0.00	0.00	0.01	0.01	0.0437	0.0172	0.0087	0.0017	0.0008
38	0	0.00	0.00	0.01	0.20	0.0514	0.0212	0.0098	0.0021	0.0010
39	0	0.00	0.00	0.01	0.50	0.0903	0.0304	0.0155	0.0033	0.0017
40	0	0.00	0.00	0.20	0.01	0.0382	0.0134	0.0078	0.0015	0.0007
41	0	0.00	0.00	0.20	0.20	0.0452	0.0167	0.0079	0.0016	0.0008
42	0	0.00	0.00	0.20	0.50	0.0636	0.0248	0.0125	0.0024	0.0013
43	0	0.00	0.00	0.50	0.01	0.0303	0.0109	0.0052	0.0011	0.0006
44	0	0.00	0.00	0.50	0.20	0.0325	0.0133	0.0060	0.0013	0.0006
45	0	0.00	0.00	0.50	0.50	0.0399	0.0173	0.0082	0.0018	0.0008
46	0	0.00	1.00	0.01	0.01	0.0284	0.0115	0.0063	0.0012	0.0006
47	0	0.00	1.00	0.01	0.20	0.0353	0.0154	0.0068	0.0015	0.0007
48	0	0.00	1.00	0.01	0.50	0.0589	0.0222	0.0113	0.0025	0.0012
49	0	0.00	1.00	0.20	0.01	0.0285	0.0105	0.0055	0.0011	0.0006
50	0	0.00	1.00	0.20	0.20	0.0340	0.0135	0.0064	0.0013	0.0006
51	0	0.00	1.00	0.20	0.50	0.0505	0.0215	0.0099	0.0020	0.0010
52	0	0.00	1.00	0.50	0.01	0.0263	0.0101	0.0053	0.0010	0.0005
53	0	0.00	1.00	0.50	0.20	0.0302	0.0119	0.0062	0.0011	0.0006
54	0	0.00	1.00	0.50	0.50	0.0449	0.0173	0.0089	0.0016	0.0008
55	0	0.50	-1.00	0.01	0.01	0.0962	0.0386	0.0194	0.0042	0.0022
56	0	0.50	-1.00	0.01	0.20	0.1296	0.0492	0.0245	0.0049	0.0030

Continued on next page

Table 4 – Unadjusted analysis (continued from previous page)

	parameter settings					empirical MSE				
setting	ϕ	ρ	ξ	p_0	p_1	n=200	n=500	n=1,000	n=5,000	n=10,000
57	0	0.50	-1.00	0.01	0.50	0.2099	0.0753	0.0403	0.0078	0.0046
58	0	0.50	-1.00	0.20	0.01	0.1043	0.0645	0.0531	0.0445	0.0439
59	0	0.50	-1.00	0.20	0.20	0.1192	0.0783	0.0675	0.0572	0.0557
60	0	0.50	-1.00	0.20	0.50	0.1817	0.1167	0.0997	0.0877	0.0864
61	0	0.50	-1.00	0.50	0.01	0.1436	0.1157	0.1078	0.1035	0.1025
62	0	0.50	-1.00	0.50	0.20	0.1620	0.1340	0.1243	0.1193	0.1176
63	0	0.50	-1.00	0.50	0.50	0.1930	0.1654	0.1614	0.1536	0.1515
64	0	0.50	0.00	0.01	0.01	0.0480	0.0197	0.0101	0.0021	0.0010
65	0	0.50	0.00	0.01	0.20	0.0651	0.0263	0.0123	0.0025	0.0012
66	0	0.50	0.00	0.01	0.50	0.1051	0.0399	0.0185	0.0038	0.0022
67	0	0.50	0.00	0.20	0.01	0.0516	0.0273	0.0188	0.0130	0.0112
68	0	0.50	0.00	0.20	0.20	0.0592	0.0348	0.0226	0.0175	0.0157
69	0	0.50	0.00	0.20	0.50	0.0920	0.0540	0.0428	0.0309	0.0301
70	0	0.50	0.00	0.50	0.01	0.0676	0.0497	0.0453	0.0402	0.0399
71	0	0.50	0.00	0.50	0.20	0.0829	0.0628	0.0538	0.0509	0.0502
72	0	0.50	0.00	0.50	0.50	0.1278	0.1021	0.0877	0.0818	0.0794
73	0	0.50	1.00	0.01	0.01	0.0310	0.0134	0.0064	0.0012	0.0006
74	0	0.50	1.00	0.01	0.20	0.0384	0.0150	0.0083	0.0015	0.0008
75	0	0.50	1.00	0.01	0.50	0.0663	0.0264	0.0125	0.0025	0.0013
76	0	0.50	1.00	0.20	0.01	0.0323	0.0133	0.0073	0.0031	0.0026
77	0	0.50	1.00	0.20	0.20	0.0373	0.0165	0.0093	0.0040	0.0036
78	0	0.50	1.00	0.20	0.50	0.0589	0.0280	0.0176	0.0086	0.0074
79	0	0.50	1.00	0.50	0.01	0.0338	0.0193	0.0150	0.0107	0.0100
80	0	0.50	1.00	0.50	0.20	0.0444	0.0237	0.0183	0.0143	0.0138
81	0	0.50	1.00	0.50	0.50	0.0698	0.0452	0.0339	0.0275	0.0272
82	1	-0.50	-1.00	0.01	0.01	0.0608	0.0232	0.0116	0.0026	0.0017
83	1	-0.50	-1.00	0.01	0.20	0.0683	0.0287	0.0141	0.0035	0.0018
84	1	-0.50	-1.00	0.01	0.50	0.1218	0.0468	0.0225	0.0062	0.0038
85	1	-0.50	-1.00	0.20	0.01	0.1497	0.1168	0.1097	0.1044	0.1042
86	1	-0.50	-1.00	0.20	0.20	0.1896	0.1559	0.1544	0.1445	0.1425
87	1	-0.50	-1.00	0.20	0.50	0.3271	0.2927	0.2786	0.2723	0.2678
88	1	-0.50	-1.00	0.50	0.01	0.3906	0.3617	0.3651	0.3573	0.3580
89	1	-0.50	-1.00	0.50	0.20	0.4861	0.4723	0.4555	0.4533	0.4517
90	1	-0.50	-1.00	0.50	0.50	0.7447	0.7186	0.7080	0.7068	0.7061
91	1	-0.50	0.00	0.01	0.01	0.0358	0.0136	0.0068	0.0014	0.0008
92	1	-0.50	0.00	0.01	0.20	0.0477	0.0175	0.0093	0.0017	0.0009
93	1	-0.50	0.00	0.01	0.50	0.0781	0.0275	0.0144	0.0030	0.0016
94	1	-0.50	0.00	0.20	0.01	0.0508	0.0322	0.0264	0.0214	0.0202
95	1	-0.50	0.00	0.20	0.20	0.0631	0.0445	0.0353	0.0294	0.0296
96	1	-0.50	0.00	0.20	0.50	0.1186	0.0817	0.0759	0.0657	0.0638
97	1	-0.50	0.00	0.50	0.01	0.1140	0.1008	0.0962	0.0921	0.0927
98	1	-0.50	0.00	0.50	0.20	0.1484	0.1372	0.1320	0.1269	0.1261
99	1	-0.50	0.00	0.50	0.50	0.2892	0.2589	0.2453	0.2413	0.2442
100	1	-0.50	1.00	0.01	0.01	0.0292	0.0114	0.0057	0.0011	0.0005
101	1	-0.50	1.00	0.01	0.20	0.0380	0.0149	0.0068	0.0016	0.0007
102	1	-0.50	1.00	0.01	0.50	0.0607	0.0234	0.0105	0.0022	0.0011
103	1	-0.50	1.00	0.20	0.01	0.0297	0.0134	0.0080	0.0043	0.0035
104	1	-0.50	1.00	0.20	0.20	0.0384	0.0178	0.0115	0.0058	0.0053
105	1	-0.50	1.00	0.20	0.50	0.0632	0.0285	0.0217	0.0131	0.0119
106	1	-0.50	1.00	0.50	0.01	0.0413	0.0261	0.0226	0.0180	0.0173
107	1	-0.50	1.00	0.50	0.20	0.0557	0.0380	0.0293	0.0263	0.0252
108	1	-0.50	1.00	0.50	0.50	0.0950	0.0711	0.0604	0.0569	0.0559
109	1	0.00	-1.00	0.01	0.01	0.0639	0.0252	0.0113	0.0025	0.0015
110	1	0.00	-1.00	0.01	0.20	0.0823	0.0288	0.0156	0.0035	0.0018
111	1	0.00	-1.00	0.01	0.50	0.1234	0.0498	0.0256	0.0054	0.0031
112	1	0.00	-1.00	0.20	0.01	0.1077	0.0798	0.0726	0.0660	0.0655
113	1	0.00	-1.00	0.20	0.20	0.1285	0.1055	0.0959	0.0889	0.0878
114	1	0.00	-1.00	0.20	0.50	0.2239	0.1853	0.1756	0.1627	0.1613
115	1	0.00	-1.00	0.50	0.01	0.2322	0.2175	0.2117	0.2094	0.2070
116	1	0.00	-1.00	0.50	0.20	0.2925	0.2672	0.2680	0.2593	0.2589
117	1	0.00	-1.00	0.50	0.50	0.4290	0.3983	0.3945	0.3874	0.3840
118	1	0.00	0.00	0.01	0.01	0.0372	0.0163	0.0084	0.0016	0.0008
119	1	0.00	0.00	0.01	0.20	0.0487	0.0186	0.0092	0.0019	0.0011
120	1	0.00	0.00	0.01	0.50	0.0825	0.0314	0.0158	0.0029	0.0017
121	1	0.00	0.00	0.20	0.01	0.0461	0.0261	0.0190	0.0142	0.0134
122	1	0.00	0.00	0.20	0.20	0.0616	0.0336	0.0264	0.0191	0.0184
123	1	0.00	0.00	0.20	0.50	0.0901	0.0636	0.0509	0.0422	0.0409

Continued on next page

Table 4 – Unadjusted analysis (continued from previous page)

setting	parameter settings ϕ	ρ	ξ	p_0	p_1	empirical MSE n=200	n=500	n=1,000	n=5,000	n=10,000
124	1	0.00	0.00	0.50	0.01	0.0819	0.0650	0.0627	0.0576	0.0572
125	1	0.00	0.00	0.50	0.20	0.1063	0.0910	0.0825	0.0791	0.0784
126	1	0.00	0.00	0.50	0.50	0.1990	0.1635	0.1551	0.1457	0.1460
127	1	0.00	1.00	0.01	0.01	0.0276	0.0118	0.0057	0.0012	0.0006
128	1	0.00	1.00	0.01	0.20	0.0370	0.0144	0.0078	0.0014	0.0008
129	1	0.00	1.00	0.01	0.50	0.0577	0.0221	0.0114	0.0023	0.0012
130	1	0.00	1.00	0.20	0.01	0.0313	0.0139	0.0074	0.0029	0.0025
131	1	0.00	1.00	0.20	0.20	0.0384	0.0171	0.0097	0.0043	0.0036
132	1	0.00	1.00	0.20	0.50	0.0632	0.0271	0.0178	0.0090	0.0084
133	1	0.00	1.00	0.50	0.01	0.0358	0.0209	0.0165	0.0123	0.0114
134	1	0.00	1.00	0.50	0.20	0.0457	0.0280	0.0230	0.0172	0.0161
135	1	0.00	1.00	0.50	0.50	0.0811	0.0525	0.0435	0.0363	0.0355
136	1	0.50	-1.00	0.01	0.01	0.0718	0.0303	0.0147	0.0030	0.0015
137	1	0.50	-1.00	0.01	0.20	0.0984	0.0340	0.0165	0.0038	0.0019
138	1	0.50	-1.00	0.01	0.50	0.1500	0.0581	0.0273	0.0059	0.0032
139	1	0.50	-1.00	0.20	0.01	0.0680	0.0419	0.0299	0.0247	0.0233
140	1	0.50	-1.00	0.20	0.20	0.0901	0.0542	0.0403	0.0324	0.0311
141	1	0.50	-1.00	0.20	0.50	0.1326	0.0806	0.0654	0.0559	0.0551
142	1	0.50	-1.00	0.50	0.01	0.0955	0.0805	0.0714	0.0671	0.0669
143	1	0.50	-1.00	0.50	0.20	0.1145	0.0962	0.0857	0.0820	0.0828
144	1	0.50	-1.00	0.50	0.50	0.1588	0.1353	0.1259	0.1167	0.1160
145	1	0.50	0.00	0.01	0.01	0.0398	0.0180	0.0080	0.0016	0.0008
146	1	0.50	0.00	0.01	0.20	0.0488	0.0215	0.0104	0.0022	0.0010
147	1	0.50	0.00	0.01	0.50	0.0893	0.0320	0.0159	0.0033	0.0018
148	1	0.50	0.00	0.20	0.01	0.0443	0.0194	0.0114	0.0058	0.0054
149	1	0.50	0.00	0.20	0.20	0.0489	0.0224	0.0161	0.0083	0.0075
150	1	0.50	0.00	0.20	0.50	0.0812	0.0401	0.0266	0.0163	0.0156
151	1	0.50	0.00	0.50	0.01	0.0464	0.0300	0.0253	0.0212	0.0207
152	1	0.50	0.00	0.50	0.20	0.0610	0.0401	0.0329	0.0274	0.0272
153	1	0.50	0.00	0.50	0.50	0.0853	0.0678	0.0580	0.0496	0.0487
154	1	0.50	1.00	0.01	0.01	0.0316	0.0117	0.0062	0.0013	0.0006
155	1	0.50	1.00	0.01	0.20	0.0376	0.0160	0.0074	0.0014	0.0008
156	1	0.50	1.00	0.01	0.50	0.0643	0.0259	0.0126	0.0025	0.0012
157	1	0.50	1.00	0.20	0.01	0.0306	0.0116	0.0065	0.0018	0.0014
158	1	0.50	1.00	0.20	0.20	0.0375	0.0164	0.0083	0.0023	0.0019
159	1	0.50	1.00	0.20	0.50	0.0547	0.0273	0.0135	0.0048	0.0038
160	1	0.50	1.00	0.50	0.01	0.0320	0.0143	0.0092	0.0052	0.0047
161	1	0.50	1.00	0.50	0.20	0.0358	0.0195	0.0121	0.0070	0.0066
162	1	0.50	1.00	0.50	0.50	0.0619	0.0283	0.0224	0.0140	0.0134

Table 5: *Empirical MSE's for an adjusted cause-specific analysis for the 162 parameter settings, based on 1000 simulations.*

setting	parameter settings ϕ	ρ	ξ	p_0	p_1	empirical MSE n=200	n=500	n=1,000	n=5,000	n=10,000
1	0	-0.50	-1.00	0.01	0.01	0.0818	0.0271	0.0137	0.0031	0.0015
2	0	-0.50	-1.00	0.01	0.20	0.0929	0.0371	0.0178	0.0033	0.0018
3	0	-0.50	-1.00	0.01	0.50	0.1701	0.0599	0.0318	0.0059	0.0031
4	0	-0.50	-1.00	0.20	0.01	0.1131	0.0452	0.0227	0.0041	0.0023
5	0	-0.50	-1.00	0.20	0.20	0.2004	0.0709	0.0328	0.0072	0.0033
6	0	-0.50	-1.00	0.20	0.50	0.6499	0.2074	0.0973	0.0203	0.0098
7	0	-0.50	-1.00	0.50	0.01	0.2734	0.0887	0.0462	0.0090	0.0046
8	0	-0.50	-1.00	0.50	0.20	0.7936	0.2404	0.1054	0.0229	0.0117
9	0	-0.50	-1.00	0.50	0.50	0.0081	0.0037	0.0024	0.0020	0.0021
10	0	-0.50	0.00	0.01	0.01	0.0386	0.0153	0.0078	0.0016	0.0008
11	0	-0.50	0.00	0.01	0.20	0.0477	0.0200	0.0094	0.0020	0.0011
12	0	-0.50	0.00	0.01	0.50	0.0883	0.0331	0.0171	0.0034	0.0017
13	0	-0.50	0.00	0.20	0.01	0.0504	0.0196	0.0096	0.0020	0.0010
14	0	-0.50	0.00	0.20	0.20	0.0724	0.0299	0.0143	0.0028	0.0013
15	0	-0.50	0.00	0.20	0.50	0.2350	0.0851	0.0453	0.0090	0.0037
16	0	-0.50	0.00	0.50	0.01	0.0754	0.0295	0.0148	0.0029	0.0015
17	0	-0.50	0.00	0.50	0.20	0.7556	0.0836	0.0430	0.0084	0.0040
18	0	-0.50	0.00	0.50	0.50	0.0032	0.0009	0.0006	0.0003	0.0003

Continued on next page

Table 5 – Adjusted analysis (continued from previous page)

	parameter settings					empirical MSE				
setting	ϕ	ρ	ξ	p_0	p_1	n=200	n=500	n=1,000	n=5,000	n=10,000
19	0	-0.50	1.00	0.01	0.01	0.0307	0.0111	0.0053	0.0011	0.0006
20	0	-0.50	1.00	0.01	0.20	0.0376	0.0141	0.0066	0.0013	0.0007
21	0	-0.50	1.00	0.01	0.50	0.0576	0.0218	0.0110	0.0023	0.0011
22	0	-0.50	1.00	0.20	0.01	0.0304	0.0118	0.0059	0.0012	0.0006
23	0	-0.50	1.00	0.20	0.20	0.0427	0.0161	0.0097	0.0016	0.0009
24	0	-0.50	1.00	0.20	0.50	0.1506	0.0565	0.0261	0.0054	0.0026
25	0	-0.50	1.00	0.50	0.01	0.0368	0.0143	0.0071	0.0016	0.0008
26	0	-0.50	1.00	0.50	0.20	0.2940	0.0450	0.0217	0.0040	0.0022
27	0	-0.50	1.00	0.50	0.50	0.0042	0.0020	0.0014	0.0012	0.0010
28	0	0.00	-1.00	0.01	0.01	0.0816	0.0318	0.0164	0.0032	0.0016
29	0	0.00	-1.00	0.01	0.20	0.1138	0.0379	0.0207	0.0039	0.0020
30	0	0.00	-1.00	0.01	0.50	0.1938	0.0776	0.0332	0.0062	0.0035
31	0	0.00	-1.00	0.20	0.01	0.1463	0.0575	0.0256	0.0056	0.0028
32	0	0.00	-1.00	0.20	0.20	0.2655	0.0964	0.0412	0.0079	0.0041
33	0	0.00	-1.00	0.20	0.50	1.0773	0.3063	0.1410	0.0270	0.0145
34	0	0.00	-1.00	0.50	0.01	0.4078	0.1445	0.0627	0.0134	0.0061
35	0	0.00	-1.00	0.50	0.20	0.9422	0.3750	0.1918	0.0323	0.0165
36	0	0.00	-1.00	0.50	0.50	0.0107	0.0063	0.0047	0.0038	0.0034
37	0	0.00	0.00	0.01	0.01	0.0446	0.0175	0.0089	0.0017	0.0008
38	0	0.00	0.00	0.01	0.20	0.0531	0.0218	0.0101	0.0022	0.0011
39	0	0.00	0.00	0.01	0.50	0.0966	0.0324	0.0165	0.0036	0.0018
40	0	0.00	0.00	0.20	0.01	0.0586	0.0199	0.0115	0.0023	0.0010
41	0	0.00	0.00	0.20	0.20	0.0883	0.0322	0.0148	0.0030	0.0016
42	0	0.00	0.00	0.20	0.50	0.4011	0.1072	0.0533	0.0097	0.0048
43	0	0.00	0.00	0.50	0.01	0.1003	0.0338	0.0174	0.0034	0.0018
44	0	0.00	0.00	0.50	0.20	0.5046	0.1091	0.0456	0.0094	0.0045
45	0	0.00	0.00	0.50	0.50	0.0097	0.0066	0.0051	0.0037	0.0033
46	0	0.00	1.00	0.01	0.01	0.0287	0.0116	0.0064	0.0012	0.0006
47	0	0.00	1.00	0.01	0.20	0.0358	0.0156	0.0069	0.0015	0.0007
48	0	0.00	1.00	0.01	0.50	0.0610	0.0230	0.0116	0.0026	0.0012
49	0	0.00	1.00	0.20	0.01	0.0335	0.0124	0.0064	0.0013	0.0006
50	0	0.00	1.00	0.20	0.20	0.0492	0.0189	0.0089	0.0019	0.0009
51	0	0.00	1.00	0.20	0.50	0.1746	0.0576	0.0260	0.0050	0.0027
52	0	0.00	1.00	0.50	0.01	0.0651	0.0170	0.0088	0.0017	0.0008
53	0	0.00	1.00	0.50	0.20	0.7482	0.0558	0.0246	0.0042	0.0022
54	0	0.00	1.00	0.50	0.50	0.0092	0.0070	0.0051	0.0038	0.0035
55	0	0.50	-1.00	0.01	0.01	0.1037	0.0416	0.0206	0.0040	0.0020
56	0	0.50	-1.00	0.01	0.20	0.1470	0.0536	0.0267	0.0051	0.0027
57	0	0.50	-1.00	0.01	0.50	0.2625	0.0881	0.0468	0.0077	0.0043
58	0	0.50	-1.00	0.20	0.01	0.2564	0.0902	0.0427	0.0072	0.0037
59	0	0.50	-1.00	0.20	0.20	0.4873	0.1473	0.0643	0.0124	0.0060
60	0	0.50	-1.00	0.20	0.50	1.5987	0.5887	0.3060	0.0391	0.0200
61	0	0.50	-1.00	0.50	0.01	0.8737	0.2634	0.1165	0.0198	0.0095
62	0	0.50	-1.00	0.50	0.20	1.4420	0.7524	0.3783	0.0591	0.0259
63	0	0.50	-1.00	0.50	0.50	0.0931	0.0901	0.0892	0.0872	0.0870
64	0	0.50	0.00	0.01	0.01	0.0495	0.0203	0.0104	0.0021	0.0010
65	0	0.50	0.00	0.01	0.20	0.0674	0.0273	0.0128	0.0025	0.0012
66	0	0.50	0.00	0.01	0.50	0.1146	0.0428	0.0199	0.0040	0.0022
67	0	0.50	0.00	0.20	0.01	0.0679	0.0279	0.0135	0.0028	0.0012
68	0	0.50	0.00	0.20	0.20	0.1074	0.0405	0.0201	0.0042	0.0019
69	0	0.50	0.00	0.20	0.50	1.5377	0.1869	0.0791	0.0122	0.0062
70	0	0.50	0.00	0.50	0.01	0.2201	0.0519	0.0241	0.0050	0.0025
71	0	0.50	0.00	0.50	0.20	0.9307	0.3155	0.1282	0.0147	0.0068
72	0	0.50	0.00	0.50	0.50	0.0620	0.0572	0.0551	0.0548	0.0536
73	0	0.50	1.00	0.01	0.01	0.0313	0.0136	0.0065	0.0013	0.0006
74	0	0.50	1.00	0.01	0.20	0.0391	0.0153	0.0085	0.0015	0.0008
75	0	0.50	1.00	0.01	0.50	0.0695	0.0275	0.0129	0.0026	0.0013
76	0	0.50	1.00	0.20	0.01	0.0384	0.0140	0.0068	0.0014	0.0007
77	0	0.50	1.00	0.20	0.20	0.0548	0.0209	0.0102	0.0021	0.0010
78	0	0.50	1.00	0.20	0.50	0.2721	0.0775	0.0357	0.0059	0.0032
79	0	0.50	1.00	0.50	0.01	0.0679	0.0192	0.0227	0.0018	0.0010
80	0	0.50	1.00	0.50	0.20	1.0008	0.1356	0.0426	0.0059	0.0028
81	0	0.50	1.00	0.50	0.50	0.0312	0.0263	0.0240	0.0230	0.0232
82	1	-0.50	-1.00	0.01	0.01	0.0639	0.0235	0.0117	0.0024	0.0012
83	1	-0.50	-1.00	0.01	0.20	0.0722	0.0300	0.0143	0.0029	0.0013
84	1	-0.50	-1.00	0.01	0.50	0.1401	0.0509	0.0230	0.0051	0.0024
85	1	-0.50	-1.00	0.20	0.01	0.0928	0.0331	0.0163	0.0030	0.0017

Continued on next page

Table 5 – Adjusted analysis (continued from previous page)

	parameter settings					empirical MSE				
setting	ϕ	ρ	ξ	p_0	p_1	n=200	n=500	n=1,000	n=5,000	n=10,000
86	1	-0.50	-1.00	0.20	0.20	0.1294	0.0532	0.0263	0.0049	0.0026
87	1	-0.50	-1.00	0.20	0.50	0.9922	0.1632	0.0719	0.0136	0.0076
88	1	-0.50	-1.00	0.50	0.01	0.1345	0.0516	0.0258	0.0062	0.0035
89	1	-0.50	-1.00	0.50	0.20	0.5570	0.1302	0.0618	0.0150	0.0105
90	1	-0.50	-1.00	0.50	0.50	0.7216	0.7267	0.7288	0.7264	0.7186
91	1	-0.50	0.00	0.01	0.01	0.0363	0.0138	0.0069	0.0014	0.0007
92	1	-0.50	0.00	0.01	0.20	0.0488	0.0180	0.0094	0.0017	0.0009
93	1	-0.50	0.00	0.01	0.50	0.0802	0.0289	0.0150	0.0029	0.0014
94	1	-0.50	0.00	0.20	0.01	0.0454	0.0184	0.0079	0.0016	0.0008
95	1	-0.50	0.00	0.20	0.20	0.0704	0.0255	0.0118	0.0025	0.0013
96	1	-0.50	0.00	0.20	0.50	0.5851	0.1737	0.0438	0.0081	0.0043
97	1	-0.50	0.00	0.50	0.01	0.4906	0.3171	0.0499	0.0028	0.0015
98	1	-0.50	0.00	0.50	0.20	2.6972	8.7270	0.0487	0.0107	0.0062
99	1	-0.50	0.00	0.50	0.50	0.4184	0.4104	0.4059	0.4023	0.3966
100	1	-0.50	1.00	0.01	0.01	0.0295	0.0115	0.0058	0.0011	0.0005
101	1	-0.50	1.00	0.01	0.20	0.0367	0.0149	0.0070	0.0016	0.0007
102	1	-0.50	1.00	0.01	0.50	0.0540	0.0222	0.0098	0.0021	0.0012
103	1	-0.50	1.00	0.20	0.01	0.0306	0.0118	0.0058	0.0012	0.0006
104	1	-0.50	1.00	0.20	0.20	0.0448	0.0179	0.0092	0.0017	0.0010
105	1	-0.50	1.00	0.20	0.50	0.9156	0.0781	0.0274	0.0046	0.0025
106	1	-0.50	1.00	0.50	0.01	2.7391	9.5637	1.7015	0.1744	0.0042
107	1	-0.50	1.00	0.50	0.20	0.5339	0.0959	0.1345	0.0546	0.1145
108	1	-0.50	1.00	0.50	0.50	0.2558	0.2624	0.2641	0.2777	0.2795
109	1	0.00	-1.00	0.01	0.01	0.0672	0.0261	0.0115	0.0025	0.0012
110	1	0.00	-1.00	0.01	0.20	0.0883	0.0300	0.0162	0.0033	0.0016
111	1	0.00	-1.00	0.01	0.50	0.1387	0.0541	0.0278	0.0048	0.0025
112	1	0.00	-1.00	0.20	0.01	0.1012	0.0372	0.0192	0.0040	0.0018
113	1	0.00	-1.00	0.20	0.20	0.1717	0.0579	0.0276	0.0054	0.0029
114	1	0.00	-1.00	0.20	0.50	0.6997	0.2053	0.0724	0.0130	0.0069
115	1	0.00	-1.00	0.50	0.01	0.2026	0.0830	0.0506	0.0943	0.1643
116	1	0.00	-1.00	0.50	0.20	0.7086	0.2491	0.0766	0.0155	0.0086
117	1	0.00	-1.00	0.50	0.50	0.5144	0.5100	0.5069	0.5059	0.5066
118	1	0.00	0.00	0.01	0.01	0.0379	0.0165	0.0085	0.0015	0.0007
119	1	0.00	0.00	0.01	0.20	0.0499	0.0190	0.0093	0.0019	0.0010
120	1	0.00	0.00	0.01	0.50	0.0869	0.0329	0.0162	0.0028	0.0016
121	1	0.00	0.00	0.20	0.01	0.0468	0.0195	0.0096	0.0017	0.0009
122	1	0.00	0.00	0.20	0.20	0.0778	0.0279	0.0138	0.0048	0.0013
123	1	0.00	0.00	0.20	0.50	0.3541	0.0926	0.0482	0.0309	0.0183
124	1	0.00	0.00	0.50	0.01	0.5680	0.4549	0.4135	0.0576	0.0017
125	1	0.00	0.00	0.50	0.20	1.2187	0.2625	0.3083	0.1367	0.2988
126	1	0.00	0.00	0.50	0.50	0.3420	0.3328	0.3285	0.3225	0.3166
127	1	0.00	1.00	0.01	0.01	0.0278	0.0119	0.0057	0.0012	0.0006
128	1	0.00	1.00	0.01	0.20	0.0370	0.0146	0.0079	0.0014	0.0008
129	1	0.00	1.00	0.01	0.50	0.0536	0.0221	0.0118	0.0023	0.0012
130	1	0.00	1.00	0.20	0.01	0.0335	0.0127	0.0063	0.0013	0.0007
131	1	0.00	1.00	0.20	0.20	0.0488	0.0181	0.0090	0.0017	0.0010
132	1	0.00	1.00	0.20	0.50	0.2477	0.1626	0.0261	0.0048	0.0026
133	1	0.00	1.00	0.50	0.01	0.5663	0.4318	0.2658	0.0978	0.0850
134	1	0.00	1.00	0.50	0.20	0.4198	0.1713	0.1326	0.0936	0.1783
135	1	0.00	1.00	0.50	0.50	0.2294	0.2330	0.2437	0.2681	0.2769
136	1	0.50	-1.00	0.01	0.01	0.0756	0.0317	0.0154	0.0031	0.0015
137	1	0.50	-1.00	0.01	0.20	0.1050	0.0362	0.0173	0.0039	0.0018
138	1	0.50	-1.00	0.01	0.50	0.1719	0.0650	0.0306	0.0060	0.0031
139	1	0.50	-1.00	0.20	0.01	0.1224	0.0460	0.0247	0.0046	0.0023
140	1	0.50	-1.00	0.20	0.20	0.2171	0.0756	0.0368	0.0068	0.0035
141	1	0.50	-1.00	0.20	0.50	0.8672	0.4305	0.1795	0.0432	0.0168
142	1	0.50	-1.00	0.50	0.01	0.3234	0.0928	0.0510	0.0118	0.0054
143	1	0.50	-1.00	0.50	0.20	1.2293	0.4280	0.2677	0.0321	0.0178
144	1	0.50	-1.00	0.50	0.50	0.3113	0.3122	0.3054	0.2936	0.2945
145	1	0.50	0.00	0.01	0.01	0.0405	0.0184	0.0081	0.0017	0.0008
146	1	0.50	0.00	0.01	0.20	0.0502	0.0221	0.0106	0.0023	0.0010
147	1	0.50	0.00	0.01	0.50	0.0950	0.0339	0.0167	0.0034	0.0019
148	1	0.50	0.00	0.20	0.01	0.0565	0.0203	0.0102	0.0020	0.0010
149	1	0.50	0.00	0.20	0.20	0.0826	0.0297	0.0166	0.0031	0.0015
150	1	0.50	0.00	0.20	0.50	0.4266	0.1180	0.0516	0.0095	0.0048
151	1	0.50	0.00	0.50	0.01	0.1441	0.0435	0.0297	0.0031	0.0016
152	1	0.50	0.00	0.50	0.20	1.2499	0.1603	0.0719	0.0208	0.0283

Continued on next page

Table 5 – Adjusted analysis (continued from previous page)

	parameter settings					empirical MSE				
setting	ϕ	ρ	ξ	p_0	p_1	n=200	n=500	n=1,000	n=5,000	n=10,000
153	1	0.50	0.00	0.50	0.50	0.2430	0.2541	0.2596	0.2791	0.2833
154	1	0.50	1.00	0.01	0.01	0.0318	0.0118	0.0062	0.0013	0.0006
155	1	0.50	1.00	0.01	0.20	0.0382	0.0162	0.0075	0.0015	0.0008
156	1	0.50	1.00	0.01	0.50	0.0632	0.0265	0.0131	0.0026	0.0013
157	1	0.50	1.00	0.20	0.01	0.0345	0.0130	0.0065	0.0013	0.0007
158	1	0.50	1.00	0.20	0.20	0.0507	0.0210	0.0101	0.0018	0.0010
159	1	0.50	1.00	0.20	0.50	0.3370	0.0766	0.0272	0.0056	0.0027
160	1	0.50	1.00	0.50	0.01	1.1240	8.1933	0.0148	0.0016	0.0009
161	1	0.50	1.00	0.50	0.20	0.8802	0.1717	0.0723	0.0043	0.0021
162	1	0.50	1.00	0.50	0.50	0.1910	0.1900	0.1901	0.1702	0.1616

2.4 A view on MSE when allowing time-varying ξ-estimation

This sections aims to illustrate what impact the estimation of a time-varying ξ can have on the MSE of the ϕ and ρ estimates. We consider one binary covariate Z influencing both the risk of the event of interest and the competing risk. We focus on a setting where our method of analysis promises some added value, i.e. $\phi = 1$, ρ=-0.5 and $p_0 = p_1 = 0.2$. As before, we set up the cause-specific hazards in such a way that the overall hazard is 0.6 in the control group. We consider both a case where the true ξ is constant and one where it is time-varying.

Each time, we perform 200 simulations at sample sizes of 100, 200, 500, 1000, 5000 (as these simulations are considerably more time-consuming, no simulations are done at a sample size of 10000). The data are always analyzed twice with our method, first assuming the true ξ is constant and next using a nonparametric estimator allowing it to vary over time.

With all these simulations, we stress that our implementation of the nonparametric estimation is probably suboptimal: little effort was spent on choosing the estimation method, and no optimization of the kernel shapes and bandwidths has been performed. This means that the reported MSE's for the nonparametric method may in fact be higher than what the method can inherently attain.

2.4.1 Constant ξ

Here, we take the true ξ constant at -1. Figure 8 compares the MSE's of both cause-specific effect estimates from the two types of analysis.

While the MSE for the effect of interest is slightly higher, the difference is small and it becomes smaller (even proportionally speaking) with increasing sample size. This means that the sacrifice that is being made in terms of efficiency is limited. Furthermore, the implicit redistribution of events that occurs in our method leads a somewhat smaller MSE for the competing risk effect in small samples.

2.4.2 Time-varying ξ

While the overall hazard in the control group is again around 0.6, we now set (under control) the cause 1-specific hazard constant at 0.161, while the cause 0-specific hazard is a sinusoidal time-varying function:

$$h_0(t) = 0.439(1 + 0.4\sin(2t))$$

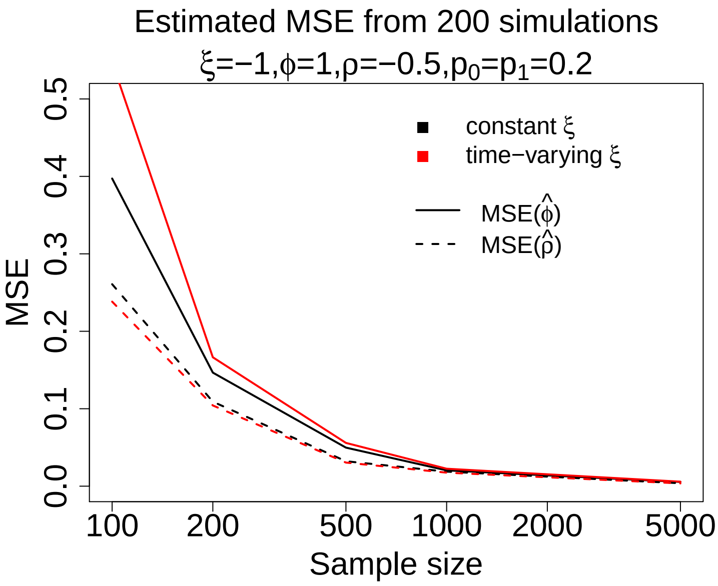


Figure 8: *Difference in MSE induced by allowing the estimation of a time-varying $\xi(t)$ in our method, assuming the true ξ is constant. The black curves show the MSE's assuming a constant ξ, the results allowing ξ to be time-varying are shown in red.*

This leads to a time-varying ξ that fluctuates around -1

$$\xi(t) = -1 - \log(1 + 0.4\sin(2t))$$

Administrative censoring occurs at $t = 9$. Figure 9 shows the MSE from these simulations.

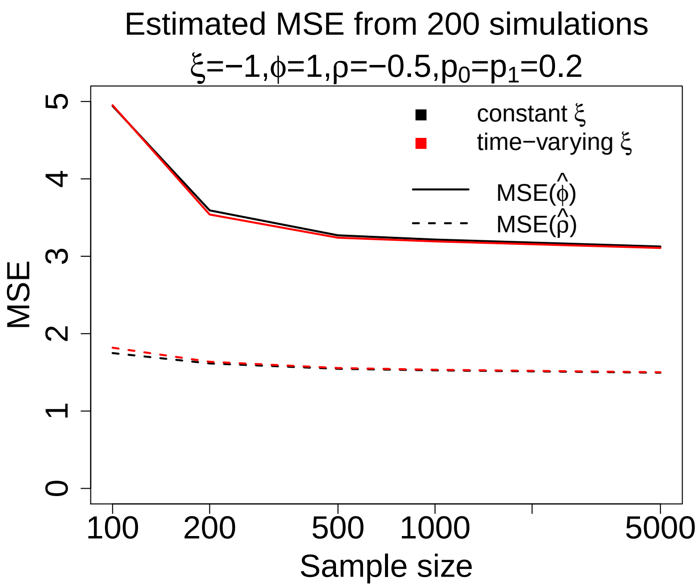


Figure 9: *Difference in MSE induced by allowing the estimation of a time-varying $\xi(t)$ in our method, assuming the true ξ is time-varying.*

This time, the difference in MSE is much smaller then before (using a truly constant ξ), even at smaller sample sizes and for both causes. Not using the nonparametric ξ-estimate hardly leads to any loss of efficiency. This may be a consequence of using only a limited time-varying effect, which also fluctuates rapidly, and neatly, around a constant value over the course of the study. This means that the effect of the time-varying ξ will effectively be balanced out over the analysis. However, simulations with less fluctuating $\xi(t)$ with bigger amplitude, namely $\xi(t) = -1 - \log(1 + 0.6 \sin((t + 2)/2.5))$ indicate that in that setting as well, little difference arises (results not shown).

3 Asymptotic covariance matrix in the time-constant ξ setting

3.1 Likelihood expressions

Under assumptions 1 and 2 from the main text and using a relative cause-specific hazard $e^{\xi(t)}$ a log partial likelihood l was constructed from the conditional probabilities of an observed event of type f at time t_i, given one such event was observed in the risk set $\mathcal{R}_i$ at that time. This likelihood distinguishes itself from the standard cause-specific partial likelihood (e.g. Fleming and Harrington[7]), by integrating contributions from both types of event, and by using the sum of cause-specific covariate-adjusted weights instead of the normal weights. These weights reflect the positive predictive value of the observed event type, conditional on the covariate values and the time, and one minus the negative predictive value.

$$
\begin{aligned}
l &= \\
&\sum_{\substack{i:C_i=1,\\F_i=0}} \left\{ \log\left(e^{-\xi(t)} e^{\boldsymbol{\rho}^T \mathbf{X}_i(t_i)} (1-p_0(t_i)) + e^{\boldsymbol{\phi}^T \mathbf{Z}_i(t_i)} p_1(t_i) \right) - \log\left(\sum_{j\in\mathcal{R}_i} \left[e^{-\xi(t)} e^{\boldsymbol{\rho}^T \mathbf{X}_j(t_i)} (1-p_0(t_i)) + e^{\boldsymbol{\phi}^T \mathbf{Z}_j(t_i)} p_1(t_i) \right] \right) \right\} \\
&+ \sum_{\substack{i:C_i=1,\\F_i=1}} \left\{ \log\left(e^{-\xi(t)} e^{\boldsymbol{\rho}^T \mathbf{X}_i(t_i)} p_0(t_i) + e^{\boldsymbol{\phi}^T \mathbf{Z}_i(t_i)} (1-p_1(t_i)) \right) - \log\left(\sum_{j\in\mathcal{R}_i} \left[e^{-\xi(t)} e^{\boldsymbol{\rho}^T \mathbf{X}_j(t_i)} p_0(t_i) + e^{\boldsymbol{\phi}^T \mathbf{Z}_j(t_i)} (1-p_1(t_i)) \right] \right) \right\}
\end{aligned}
$$

As pointed out in Van Rompaye, Jaffar and Goetghebeur,[6] since l follows from conditioning on the event type, all information concerning the contrast between event types and hence $\xi(t)$ is lost. If $\xi(t)$ is unknown in advance, it can be estimated by conditioning on the occurrence of any type of event in the risk set as in Dewanji.[8] This leads to a second log partial likelihood l^*:

$$
\begin{aligned}
l^* &= \sum_{\substack{i:C_i=1,\\F_i=0}} \left[\log\left(e^{-\xi(t)} e^{\boldsymbol{\rho}^T \mathbf{X}_i(t)} (1-p_0(t)) + e^{\boldsymbol{\phi}^T \mathbf{Z}_i(t)} p_1(t) \right) - \log\left(\sum_{j\in\mathcal{R}_i} \left[e^{-\xi(t)} e^{\boldsymbol{\rho}^T \mathbf{X}_i(t)} + e^{\boldsymbol{\phi}^T \mathbf{Z}_i(t)} \right] \right) \right] \\
&+ \sum_{\substack{i:C_i=1,\\F_i=1}} \left[\log\left(e^{-\xi(t)} e^{\boldsymbol{\rho}^T \mathbf{X}_i(t)} p_0(t) + e^{\boldsymbol{\phi}^T \mathbf{Z}_i(t)} (1-p_1(t)) \right) - \log\left(\sum_{j\in\mathcal{R}_i} \left[e^{-\xi(t)} e^{\boldsymbol{\rho}^T \mathbf{X}_i(t)} + e^{\boldsymbol{\phi}^T \mathbf{Z}_i(t)} \right] \right) \right]
\end{aligned}
$$

Under proportional cause-specific baseline hazards (assumption 3' from the main text, thus replacing $\xi(t)$ by ξ throughout), this becomes a finite dimensional parametric log likelihood.

3.2 Outline of the derivation of the asymptotic covariance matrix for the estimates from $\mathbf{T}(\boldsymbol{\theta}) = 0$

We call the full set of parameters $\boldsymbol{\theta} = \{\xi, \boldsymbol{\phi}, \boldsymbol{\rho}\}$ and the vector of true parameter values $\boldsymbol{\theta}^0$. The left hand side of the joint estimating equations under assumption 3'

$$\mathbf{T}(\boldsymbol{\theta}) = \left(\frac{\partial l}{\partial \boldsymbol{\phi}}, \frac{\partial l}{\partial \boldsymbol{\rho}}, \frac{\partial l^*}{\partial \xi} \right) = 0 \tag{3.1}$$

is a vector of martingales, which can be seen by inserting the Doob-Meyer decomposition of the various counting processes.[3] Using counting process theory one can show consistency and asymptotic normality of the solutions to these equations. This is based on first showing asymptotic normality for the components of $\mathbf{T}(\boldsymbol{\theta}^0)$, and then linking these to the components of $\mathbf{T}(\boldsymbol{\theta})$ through a Taylor series expansion (as in Andersen and Gill[4]). In doing this, using two different likelihoods l and l^* leads to a complication handled in the same way as in Goetghebeur and Ryan.[5] By applying Slutsky's theorem to the Taylor expansion one can obtain a consistent estimate for the asymptotic covariance matrix from a consistent estimator for the covariance process under the null. Asymptotically, we get:

$$n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) \sim N(0, \boldsymbol{\Sigma}) = N(0, \boldsymbol{\Gamma}^{-1} \boldsymbol{\Delta} \boldsymbol{\Gamma}^{-1T}) \tag{3.2}$$

where $\boldsymbol{\Delta}$ denotes the asymptotic covariance matrix of $n^{-1/2}\mathbf{T}(\boldsymbol{\theta}^0)$ (from the predictable variation process) and $\boldsymbol{\Gamma}$ is the limit of $1/n$ times the matrix of derivatives of $\mathbf{T}$ in $\boldsymbol{\theta}^0$. We now derive the exact expressions of the components in this sandwich variance structure.

3.3 Expression for the asymptotic covariance matrix for the estimates from $\mathbf{T}(\boldsymbol{\theta}) = 0$

In this section, we will use counting process notation, referring to the individual processes representing the observed failures:

$$\begin{aligned} N_{i0}(t) &= I[D_i \le t, C_i = 1, F_i = 0] \\ N_{i1}(t) &= I[D_i \le t, C_i = 1, F_i = 1] \\ N_{i\bullet}(t) &= N_{i0}(t) + N_{i1}(t) \end{aligned}$$

The at-risk of observation indicator for person i is called $Y_i(t) \in \{0, 1\}$, the maximal follow-up time is called τ.

We will also apply the following shorthand notation:

$$\begin{aligned} w_i^0(t) &= e^{-\xi} e^{\boldsymbol{\rho}^T \mathbf{X}_i(t)} \\ w_i^1(t) &= e^{\boldsymbol{\phi}^T \mathbf{Z}_i(t)} \end{aligned}$$

Before we give the expressions for Γ, we first remind the reader of the components of $\mathbf{T}$:

$$\frac{\partial l}{\partial \boldsymbol{\phi}} = \sum_{i=1}^{n}\int_0^{\tau}\left[\frac{\mathbf{Z}_i(t)w_i^1(t)p_1(t)}{w_i^0(t)(1-p_0(t))+w_i^1(t)p_1(t)} - \frac{\sum_{j=1}^{n} Y_j(t)\mathbf{Z}_j(t)w_j^1(t)p_1(t)}{\sum_{j=1}^{n} Y_j(t)\left[w_j^0(t)(1-p_0(t))+w_j^1(t)p_1(t)\right]}\right]dN_{i0}(t)$$
$$+\sum_{i=1}^{n}\int_0^{\tau}\left[\frac{\mathbf{Z}_i(t)w_i^1(t)(1-p_1(t))}{w_i^0(t)p_0(t)+w_i^1(t)(1-p_1(t))} - \frac{\sum_{j=1}^{n} Y_j(t)\mathbf{Z}_j(t)w_j^1(t)(1-p_1(t))}{\sum_{j=1}^{n} Y_j(t)\left[w_j^0(t)p_0(t)+w_j^1(t)(1-p_1(t))\right]}\right]dN_{i1}(t)$$

$$\frac{\partial l}{\partial \boldsymbol{\rho}} = \sum_{i=1}^{n}\int_0^{\tau}\left[\frac{\mathbf{X}_i(t)w_i^0(t)(1-p_0(t))}{w_i^0(t)(1-p_0(t))+w_i^1(t)p_1(t)} - \frac{\sum_{j=1}^{n} Y_j(t)\mathbf{X}_j(t)w_j^0(t)(1-p_0(t))}{\sum_{j=1}^{n} Y_j(t)\left[w_j^0(t)(1-p_0(t))+w_j^1(t)p_1(t)\right]}\right]dN_{i0}(t)$$
$$+\sum_{i=1}^{n}\int_0^{\tau}\left[\frac{\mathbf{X}_i(t)w_i^0(t)p_0(t)}{w_i^0(t)p_0(t)+w_i^1(t)(1-p_1(t))} - \frac{\sum_{j=1}^{n} Y_j(t)\left[\mathbf{X}_j(t)w_j^0(t)p_0(t)\right]}{\sum_{j=1}^{n} Y_j(t)\left[w_j^0(t)p_0(t)+w_j^1(t)(1-p_1(t))\right]}\right]dN_{i1}(t)$$

$$\frac{\partial l^*}{\partial \xi} = \sum_{i=1}^{n}\int_0^{\tau}\left[\frac{-w_i^0(t)(1-p_0(t))}{w_i^0(t)(1-p_0(t))+w_i^1(t)p_1(t)} - \frac{\sum_{j=1}^{n} -Y_j(t)w_j^0(t)}{\sum_{j=1}^{n} Y_j(t)\left[w_j^0(t)+w_j^1(t)\right]}\right]dN_{i0}(t)$$
$$+\sum_{i=1}^{n}\int_0^{\tau}\left[\frac{-w_i^0(t)p_0(t)}{\left[w_i^0(t)p_0(t)+w_i^1(t)(1-p_1(t))\right]} - \frac{\sum_{j=1}^{n} -Y_j(t)w_j^0(t)}{\sum_{j=1}^{n} Y_j(t)\left[w_j^0(t)+w_j^1(t)\right]}\right]dN_{i1}(t)$$

The various components of Γ now are:

$$\frac{\partial^2 l}{\partial \boldsymbol{\phi}^2} =$$
$$\sum_{i=1}^{n}\int_0^{\tau}\left[\frac{\mathbf{Z}_i^2(t)w_i^1(t)p_1(t)w_i^0(t)(1-p_0(t))}{\left(w_i^0(t)(1-p_0(t))+w_i^1(t)p_1(t)\right)^2}\right.$$
$$\left.-\frac{\sum_{j=1}^{n} Y_j(t)\mathbf{Z}_j^2(t)w_j^1(t)p_1(t)}{\sum_{j=1}^{n} Y_j(t)\left[w_j^0(t)(1-p_0(t))+w_j^1(t)p_1(t)\right]} + \frac{\left(\sum_{j=1}^{n} Y_j(t)\mathbf{Z}_j(t)w_j^1(t)p_1(t)\right)^2}{\left(\sum_{j=1}^{n} Y_j(t)\left[w_j^0(t)(1-p_0(t))+w_j^1(t)p_1(t)\right]\right)^2}\right]dN_{i0}(t)$$
$$+\sum_{i=1}^{n}\int_0^{\tau}\left[\frac{\mathbf{Z}_i^2(t)w_i^1(t)(1-p_1(t))w_i^0(t)p_0(t)}{\left(w_i^0(t)p_0(t)+w_i^1(t)(1-p_1(t))\right)^2}\right.$$
$$\left.-\frac{\sum_{j=1}^{n} Y_j(t)\mathbf{Z}_j^2(t)w_j^1(t)(1-p_1(t))}{\sum_{j=1}^{n} Y_j(t)\left[w_j^0(t)p_0(t)+w_j^1(t)(1-p_1(t))\right]} + \frac{\left(\sum_{j=1}^{n} Y_j(t)\mathbf{Z}_j(t)w_j^1(t)(1-p_1(t))\right)^2}{\left(\sum_{j=1}^{n} Y_j(t)\left[w_j^0(t)p_0(t)+w_j^1(t)(1-p_1(t))\right]\right)^2}\right]dN_{i1}(t)$$

The second derivative with respect to $\boldsymbol{\rho}$ is almost identical in form:

$$\frac{\partial^2 l}{\partial \boldsymbol{\rho}^2} =$$

$$\sum_{i=1}^{n}\int_0^{\tau}\left[\frac{\mathbf{X}_i^2(t)w_i^0(t)(1-p_0(t))w_i^1(t)p_1(t)}{\left(w_i^0(t)(1-p_0(t))+w_i^1(t)p_1(t)\right)^2}\right.$$

$$\left.-\frac{\sum_{j=1}^{n}Y_j(t)\mathbf{X}_j^2(t)w_j^0(t)(1-p_0(t))}{\sum_{j=1}^{n}Y_j(t)\left[w_j^0(t)(1-p_0(t))+w_j^1(t)p_1(t)\right]}+\frac{\left(\sum_{j=1}^{n}Y_j(t)\mathbf{X}_j(t)w_j^0(t)(1-p_0(t))\right)^2}{\left(\sum_{j=1}^{n}Y_j(t)\left[w_j^0(t)(1-p_0(t))+w_j^1(t)p_1(t)\right]\right)^2}\right]dN_{i0}(t)$$

$$+\sum_{i=1}^{n}\int_0^{\tau}\left[\frac{\mathbf{X}_i^2(t)w_i^0(t)p_0(t)w_i^1(t)(1-p_0(t))}{\left(w_i^0(t)p_0(t)+w_i^1(t)(1-p_1(t))\right)^2}\right.$$

$$\left.-\frac{\sum_{j=1}^{n}Y_j(t)\mathbf{X}_j^2(t)w_j^0(t)p_0(t)}{\sum_{j=1}^{n}Y_j(t)\left[w_j^0(t)p_0(t)+w_j^1(t)(1-p_1(t))\right]}+\frac{\left(\sum_{j=1}^{n}Y_j(t)\mathbf{X}_j(t)w_j^0(t)p_0(t)\right)^2}{\left(\sum_{j=1}^{n}Y_j(t)\left[w_j^0(t)p_0(t)+w_j^1(t)(1-p_1(t))\right]\right)^2}\right]dN_{i1}(t)$$

For the second derivative with respect to ξ:

$$\begin{aligned}\frac{\partial^2 l^*}{\partial \xi^2} &= \sum_{i=1}^{n}\int_0^{\tau}\left[\frac{w_i^0(t)(1-p_0(t))w_i^1(t)p_1(t)}{\left\{w_i^0(t)(1-p_0(t))+w_i^1(t)p_1(t)\right\}^2}-\frac{\sum_{j=1}^{n}Y_j(t)w_j^0(t)\sum_{j=1}^{n}Y_j(t)w_j^1(t)}{\left(\sum_{j=1}^{n}Y_j(t)\left[w_j^0(t)+w_j^1(t)\right]\right)^2}\right]dN_{i0}(t)\\ &\quad+\sum_{i=1}^{n}\int_0^{\tau}\left[\frac{w_i^0(t)p_0(t)w_i^1(t)(1-p_1(t))}{\left\{w_i^0(t)p_0(t)+w_i^1(t)(1-p_1(t))\right\}^2}-\frac{\sum_{j=1}^{n}Y_j(t)w_j^0(t)\sum_{j=1}^{n}Y_j(t)w_j^1(t)}{\left(\sum_{j=1}^{n}Y_j(t)\left[w_j^0(t)+w_j^1(t)\right]\right)^2}\right]dN_{i1}(t)\end{aligned}$$

For the second derivative of l^* with respect to ξ and $\boldsymbol{\rho}$:

$$\begin{aligned}\frac{\partial^2 l^*}{\partial \boldsymbol{\rho}\partial \xi} &= \sum_{i=1}^{n}\int_0^{\tau}\left[-\frac{w_i^0(t)(1-p_0(t))w_i^1(t)p_1(t)X_i(t)}{\left\{w_i^0(t)(1-p_0(t))+w_i^1(t)p_1(t)\right\}^2}+\frac{\sum_{j=1}^{n}Y_j(t)w_j^0(t)X_j(t)\sum_{j=1}^{n}Y_j(t)w_j^1(t)}{\left(\sum_{j=1}^{n}Y_j(t)\left[w_j^0(t)+w_j^1(t)\right]\right)^2}\right]dN_{i0}(t)\\ &\quad+\sum_{i=1}^{n}\int_0^{\tau}\left[-\frac{w_i^0(t)p_0(t)w_i^1(t)(1-p_1(t))X_i(t)}{\left\{w_i^0(t)p_0(t)+w_i^1(t)(1-p_1(t))\right\}^2}+\frac{\sum_{j=1}^{n}Y_j(t)w_j^0(t)X_j(t)\sum_{j=1}^{n}Y_j(t)w_j^1(t)}{\left(\sum_{j=1}^{n}Y_j(t)\left[w_j^0(t)+w_j^1(t)\right]\right)^2}\right]dN_{i1}(t)\end{aligned}$$

The second order derivative of l^* w.r.t. $\boldsymbol{\phi}$ and ξ:

$$\begin{aligned}\frac{\partial^2 l^*}{\partial \boldsymbol{\phi}\partial \xi} &= \sum_{i=1}^{n}\int_0^{\tau}\left[\frac{Z_i(t)w_i^0(t)(1-p_0(t))w_i^1(t)p_1(t)}{\left\{w_i^0(t)(1-p_0(t))+w_i^1(t)p_1(t)\right\}^2}-\frac{\sum_{j=1}^{n}Y_j(t)w_j^0(t)\sum_{j=1}^{n}Y_j(t)Z_j(t)w_j^1(t)}{\left(\sum_{j=1}^{n}Y_j(t)\left[w_j^0(t)+w_j^1(t)\right]\right)^2}\right]dN_{i0}(t)\\ &\quad+\sum_{i=1}^{n}\int_0^{\tau}\left[\frac{Z_i(t)w_i^0(t)p_0(t)w_i^1(t)(1-p_1(t))}{\left\{w_i^0(t)p_0(t)+w_i^1(t)(1-p_1(t))\right\}^2}-\frac{\sum_{j=1}^{n}Y_j(t)w_j^0(t)\sum_{j=1}^{n}Y_j(t)Z_j(t)w_j^1(t)}{\left(\sum_{j=1}^{n}Y_j(t)\left[w_j^0(t)+w_j^1(t)\right]\right)^2}\right]dN_{i0}(t)\end{aligned}$$

The second order derivative w.r.t. $\boldsymbol{\phi}$ and $\boldsymbol{\rho}$ (which is the same as deriving w.r.t. $\boldsymbol{\rho}$ and $\boldsymbol{\phi}$):

$$\frac{\partial^2 l}{\partial \boldsymbol{\rho} \partial \boldsymbol{\phi}} =$$

$$\sum_{i=1}^{n} \int_0^{\tau} \left[\frac{-Z_i(t)X_i(t)w_i^0(t)(1-p_0(t))w_i^1(t)p_1(t)}{\left\{w_i^0(t)(1-p_0(t))+w_i^1(t)p_1(t)\right\}^2} + \frac{\sum\limits_{j=1}^{n} Y_j(t)X_j(t)w_j^0(t)(1-p_0(t)) \sum\limits_{j=1}^{n} Y_j(t)Z_j(t)w_j^1(t)p_1(t)}{\left(\sum\limits_{j=1}^{n} Y_j(t)\left[w_j^0(t)(1-p_0(t))+w_j^1(t)p_1(t)\right]\right)^2} \right] dN_{i0}(t)$$

$$+\sum_{i=1}^{n} \int_0^{\tau} \left[\frac{-Z_i(t)X_i(t)w_i^0(t)p_0(t)w_i^1(t)(1-p_1(t))}{\left\{w_i^0(t)p_0(t)+w_i^1(t)(1-p_1(t))\right\}^2} + \frac{\sum\limits_{j=1}^{n} Y_j(t)X_j(t)w_j^0(t)p_0(t) \sum\limits_{j=1}^{n} Y_j(t)Z_j(t)w_j^1(t)(1-p_1(t))}{\left(\sum\limits_{j=1}^{n} Y_j(t)\left[w_j^0(t)p_0(t)+w_j^1(t)(1-p_1(t))\right]\right)^2} \right] dN_{i0}(t)$$

The second order derivative of l w.r.t. ξ and $\boldsymbol{\phi}$:

$$\frac{\partial^2 l}{\partial \boldsymbol{\phi} \partial \xi} =$$

$$\sum_{i=1}^{n} \int_0^{\tau} \left[\frac{Z_i(t)w_i^0(t)(1-p_0(t))w_i^1(t)p_1(t)}{\left\{w_i^0(t)(1-p_0(t))+w_i^1(t)p_1(t)\right\}^2} - \frac{\sum\limits_{j=1}^{n} Y_j(t)w_j^0(t)(1-p_0(t)) \sum\limits_{j=1}^{n} Y_j(t)Z_j(t)w_j^1(t)p_1(t)}{\left(\sum\limits_{j=1}^{n} Y_j(t)\left[w_j^0(t)(1-p_0(t))+w_j^1(t)p_1(t)\right]\right)^2} \right] dN_{i0}(t)$$

$$+\sum_{i=1}^{n} \int_0^{\tau} \left[\frac{Z_i(t)w_i^0(t)p_0(t)w_i^1(t)(1-p_1(t))}{\left\{w_i^0(t)p_0(t)+w_i^1(t)(1-p_1(t))\right\}^2} - \frac{\sum\limits_{j=1}^{n} Y_j(t)w_j^0(t)p_0(t) \sum\limits_{j=1}^{n} Y_j(t)Z_j(t)w_j^1(t)(1-p_1(t))}{\left(\sum\limits_{j=1}^{n} Y_j(t)\left[w_j^0(t)p_0(t)+w_j^1(t)(1-p_1(t))\right]\right)^2} \right] dN_{i0}(t)$$

And finally the second order derivative w.r.t. ξ and $\boldsymbol{\rho}$:

$$\frac{\partial^2 l}{\partial \boldsymbol{\rho} \partial \xi} =$$

$$\sum_{i=1}^{n} \int_0^{\tau} \left[\frac{-X_i(t)w_i^0(t)(1-p_0(t))w_i^1(t)p_1(t)}{\left\{w_i^0(t)(1-p_0(t))+w_i^1(t)p_1(t)\right\}^2} + \frac{\sum\limits_{j=1}^{n} Y_j(t)X_j(t)w_j^0(t)(1-p_0(t)) \sum\limits_{j=1}^{n} Y_j(t)w_j^1(t)p_1(t)}{\left(\sum\limits_{j=1}^{n} Y_j(t)\left[w_j^0(t)(1-p_0(t))+w_j^1(t)p_1(t)\right]\right)^2} \right] dN_{i0}(t)$$

$$+\sum_{i=1}^{n} \int_0^{\tau} \left[\frac{-X_i(t)w_i^0(t)p_0(t)w_i^1(t)(1-p_1(t))}{\left\{w_i^0(t)p_0(t)+w_i^1(t)(1-p_1(t))\right\}^2} + \frac{\sum\limits_{j=1}^{n} Y_j(t)X_j(t)w_j^0(t)p_0(t) \sum\limits_{j=1}^{n} Y_j(t)w_j^1(t)(1-p_1(t))}{\left(\sum\limits_{j=1}^{n} Y_j(t)\left[w_j^0(t)p_0(t)+w_j^1(t)(1-p_1(t))\right]\right)^2} \right] dN_{i0}(t)$$

To estimate Γ it suffices to replace the components of $\boldsymbol{\theta}$ throughout by their estimators $\hat{\boldsymbol{\theta}}$.

Next we define the estimators of the components of the symmetrical matrix $\boldsymbol{\Delta}$, the matrix of covariation processes of the martingale processes contained in $\mathbf{T}$:

$$\boldsymbol{\Delta}_{\boldsymbol{\phi},\boldsymbol{\phi}} = \sum_{i=1}^{n}\int_0^{\tau}\left[\frac{\left[\mathbf{Z}_i(t)w_i^1(t)p_1(t)\right]^2}{\left[w_i^0(t)(1-p_0(t))+w_i^1(t)p_1(t)\right]^2} - \frac{\left(\sum_{j=1}^{n}Y_j(t)\mathbf{Z}_j(t)w_j^1(t)p_1(t)\right)^2}{\left(\sum_{j=1}^{n}Y_j(t)\left[w_j^0(t)(1-p_0(t))+w_j^1(t)p_1(t)\right]\right)^2}\right]dN_{i0}(t)$$

$$+\sum_{i=1}^{n}\int_0^{\tau}\left[\frac{\left[\mathbf{Z}_i(t)w_i^1(t)(1-p_1(t))\right]^2}{\left[w_i^0(t)p_0(t)+w_i^1(t)(1-p_1(t))\right]^2} - \frac{\left(\sum_{j=1}^{n}Y_j(t)\mathbf{Z}_j(t)w_j^1(t)(1-p_1(t))\right)^2}{\left(\sum_{j=1}^{n}Y_j(t)\left[w_j^0(t)p_0(t)+w_j^1(t)(1-p_1(t))\right]\right)^2}\right]dN_{i1}(t)$$

$$\boldsymbol{\Delta}_{\boldsymbol{\rho},\boldsymbol{\rho}} = \sum_{i=1}^{n}\int_0^{\tau}\left[\left(\frac{\mathbf{X}_i(t)w_i^0(t)(1-p_0(t))}{w_i^0(t)(1-p_0(t))+w_i^1(t)p_1(t)}\right)^2 - \left(\frac{\sum_{j=1}^{n}Y_j(t)\mathbf{X}_j(t)w_j^0(t)(1-p_0(t))}{\sum_{j=1}^{n}Y_j(t)\left[w_j^0(t)(1-p_0(t))+w_j^1(t)p_1(t)\right]}\right)^2\right]dN_{i0}(t)$$

$$+\sum_{i=1}^{n}\int_0^{\tau}\left[\left(\frac{\mathbf{X}_i(t)w_i^0(t)p_0(t)}{w_i^0(t)p_0(t)+w_i^1(t)(1-p_1(t))}\right)^2 - \left(\frac{\sum_{j=1}^{n}Y_j(t)\mathbf{X}_j(t)w_j^0(t)p_0(t)}{\sum_{j=1}^{n}Y_j(t)\left[w_j^0(t)p_0(t)+w_j^1(t)(1-p_1(t))\right]}\right)^2\right]dN_{i1}(t)$$

$$\boldsymbol{\Delta}_{\xi,\xi} = \sum_{i=1}^{n}\int_0^{\tau}\left(\frac{w_i^0(t)(1-p_0(t))}{w_i^0(t)(1-p_0(t))+w_i^1(t)p_1(t)}\right)^2 dN_{i0}(t) + \sum_{i=1}^{n}\int_0^{\tau}\left(\frac{w_i^0(t)p_0(t)}{w_i^0(t)p_0(t)+w_i^1(t)(1-p_1(t))}\right)^2 dN_{i1}(t)$$

$$-\sum_{i=1}^{n}\int_0^{\tau}\left(\frac{\sum_{j=1}^{n}Y_j(t)w_j^0(t)}{\sum_{j=1}^{n}Y_j(t)\left[w_j^0(t)+w_j^1(t)\right]}\right)^2 dN_{i\bullet}(t)$$

$$\boldsymbol{\Delta}_{\boldsymbol{\phi},\boldsymbol{\rho}} = \boldsymbol{\Delta}_{\boldsymbol{\rho},\boldsymbol{\phi}} =$$

$$\sum_{i=1}^{n}\int_0^{\tau}\left[\frac{\mathbf{Z}_i(t)\mathbf{X}_i(t)w_i^1(t)w_i^0(t)p_1(t)(1-p_0(t))}{\left[w_i^0(t)(1-p_0(t))+w_i^1(t)p_1(t)\right]^2} - \frac{\sum_{j=1}^{n}Y_j(t)\mathbf{Z}_j(t)w_j^1(t)p_1(t)\sum_{j=1}^{n}Y_j(t)\mathbf{X}_j(t)w_j^0(t)(1-p_0(t))}{\left(\sum_{j=1}^{n}Y_j(t)\left[w_j^0(t)(1-p_0(t))+w_j^1(t)p_1(t)\right]\right)^2}\right]dN_{i0}(t)$$

$$+\sum_{i=1}^{n}\int_0^{\tau}\left[\frac{\mathbf{Z}_i(t)\mathbf{X}_i(t)w_i^1(t)w_i^0(t)p_0(t)(1-p_1(t))}{\left[w_i^0(t)p_0(t)+w_i^1(t)(1-p_1(t))\right]^2} - \frac{\sum_{j=1}^{n}Y_j(t)\mathbf{Z}_j(t)w_j^1(t)(1-p_1(t))\sum_{j=1}^{n}Y_j(t)\mathbf{X}_j(t)w_j^0(t)p_0(t)}{\left(\sum_{j=1}^{n}Y_j(t)\left[w_j^0(t)p_0(t)+w_j^1(t)(1-p_1(t))\right]\right)^2}\right]dN_{i1}(t)$$

$$\boldsymbol{\Delta}_{\boldsymbol{\phi},\xi} = \boldsymbol{\Delta}_{\xi,\boldsymbol{\phi}} =$$

$$\sum_{i=1}^{n}\int_0^{\tau}\left[-\frac{\mathbf{Z}_i(t)w_i^1(t)p_1(t)w_i^0(t)(1-p_0(t))}{\left[w_i^0(t)(1-p_0(t))+w_i^1(t)p_1(t)\right]^2} + \frac{\sum_{j=1}^{n}Y_j(t)\mathbf{Z}_j(t)w_j^1(t)p_1(t)\sum_{j=1}^{n}Y_j(t)w_j^0(t)(1-p_0(t))}{\left(\sum_{j=1}^{n}Y_j(t)\left[w_j^0(t)(1-p_0(t))+w_j^1(t)p_1(t)\right]\right)^2}\right]dN_{i0}(t)$$

$$+\sum_{i=1}^{n}\int_0^{\tau}\left[-\frac{\mathbf{Z}_i(t)w_i^1(t)(1-p_1(t))w_i^0(t)p_0(t)}{\left[w_i^0(t)p_0(t)+w_i^1(t)(1-p_1(t))\right]^2} + \frac{\sum_{j=1}^{n}Y_j(t)\mathbf{Z}_j(t)w_j^1(t)(1-p_1(t))\sum_{j=1}^{n}Y_j(t)w_j^0(t)p_0(t)}{\left(\sum_{j=1}^{n}Y_j(t)\left[w_j^0(t)p_0(t)+w_j^1(t)(1-p_1(t))\right]\right)^2}\right]dN_{i1}(t)$$

$$\boldsymbol{\Delta}_{\boldsymbol{\rho},\xi} = \boldsymbol{\Delta}_{\xi,\boldsymbol{\rho}} =$$

$$\sum_{i=1}^{n}\int_0^{\tau}\left[-\frac{\mathbf{X}_i(t)\left(w_i^0(t)\right)^2(1-p_0(t))^2}{\left[w_i^0(t)(1-p_0(t))+w_i^1(t)p_1(t)\right]^2} + \frac{\sum_{j=1}^{n}Y_j(t)\mathbf{X}_j(t)w_j^0(t)(1-p_0(t))\sum_{j=1}^{n}Y_j(t)w_j^0(t)(1-p_0(t))}{\left(\sum_{j=1}^{n}Y_j(t)\left[w_j^0(t)(1-p_0(t))+w_j^1(t)p_1(t)\right]\right)^2}\right]dN_{i0}(t)$$

$$+\sum_{i=1}^{n}\int_0^{\tau}\left[-\frac{\mathbf{X}_i(t)\left(w_i^0(t)\right)^2 p_0(t)^2}{\left[w_i^0(t)p_0(t)+w_i^1(t)(1-p_1(t))\right]^2} + \frac{\sum_{j=1}^{n}Y_j(t)\mathbf{X}_j(t)w_j^0(t)p_0(t)\sum_{j=1}^{n}Y_j(t)w_j^0(t)p_0(t)}{\left(\sum_{j=1}^{n}Y_j(t)\left[w_j^0(t)p_0(t)+w_j^1(t)(1-p_1(t))\right]\right)^2}\right]dN_{i1}(t)$$

4 Breslow-type estimator for the cause-specific cumulative hazards

We again use standard counting process notation, and indicate use $N_{i\bullet}(t)$ to indicate the all-cause counting process for person i. For the intensity process of the all-cause counting process, we have:

$$\lambda_{i\bullet}(t) \;=\; Y_i(t)h_1(t)\left[e^{-\xi(t)}e^{\rho^T\mathbf{X}_i(t)} + e^{\phi^T\mathbf{Z}_i(t)}\right]$$

The Doob-Meyer decomposition is:

$$dN_{i\bullet}(t) \;=\; Y_i(t)h_1(t)\left[e^{-\xi(t)}e^{\rho^T\mathbf{X}_i(t)} + e^{\phi^T\mathbf{Z}_i(t)}\right]dt + dM_{i\bullet}(t)$$

or for the counting process for all individuals combined:

$$dN_{\bullet}(t) \;=\; \sum_{i=1}^{n} Y_i(t)h_1(t)\left[e^{-\xi(t)}e^{\rho^T\mathbf{X}_i(t)} + e^{\phi^T\mathbf{Z}_i(t)}\right]dt + dM_{\bullet}(t)$$

We now introduce $J(t)$ which is zero when all $Y_i(t)$ are zero, and heuristically assume that 0/0=0 to avoid problems in our next step.

$$\frac{J(t)dN_{\bullet}(t)}{\sum_{i=1}^{n} Y_i(t)\left[e^{-\xi(t)}e^{\rho^T\mathbf{X}_i(t)} + e^{\phi^T\mathbf{Z}_i(t)}\right]} \;=\; J(t)h_1(t)dt + \frac{J(t)dM_{\bullet}(t)}{\sum_{i=1}^{n} Y_i(t)\left[e^{-\xi(t)}e^{\rho^T\mathbf{X}_i(t)} + e^{\phi^T\mathbf{Z}_i(t)}\right]}$$

and integrating over time:

$$\int_0^t \frac{J(s)dN_{\bullet}(s)}{\sum_{i=1}^{n} Y_i(s)\left[e^{-\xi(s)}e^{\rho^T\mathbf{X}_i(s)} + e^{\phi^T\mathbf{Z}_i(s)}\right]} \;=\; \int_0^t J(s)h_1(s)ds + \int_0^t \frac{J(s)dM_{\bullet}(s)}{\sum_{i=1}^{n} Y_i(s)\left[e^{-\xi(s)}e^{\rho^T\mathbf{X}_i(s)} + e^{\phi^T\mathbf{Z}_i(s)}\right]}$$

Since the last term is a stochastic integral w.r.t. a martingale, it is itself a mean zero martingale. This means that the left-hand side is in fact an unbiased estimator of the first term on the right-hand side, which is one of the cumulative hazards we are aiming for (apart from the slight complication of $J(s)$, which tells us we cannot estimate anything if there are no patients at risk). To evaluate this expression we replace ξ, ϕ and ρ by their estimators. Our estimator of the baseline cumulative type 1 hazard thus becomes:

$$\hat{\Lambda}_0^1(t) \;=\; \int_0^t \frac{1}{\sum_{i=1}^{n} Y_i(s)\left[e^{-\hat{\xi}(s)}e^{\hat{\rho}^T\mathbf{X}_i(s)} + e^{\hat{\phi}^T\mathbf{Z}_i(s)}\right]}\left(\sum_{i=1}^{n} dN_{\bullet}(s)\right)$$

In the absence of misclassification, this expression reduces to the standard Breslow estimator.

To estimate the baseline hazard for the competing risks we proceed along the same line and get:

$$\hat{\Lambda}_0^0(t) \;=\; \int_0^t \frac{1}{\sum_{i=1}^{n} Y_i(s)\left[e^{\hat{\rho}^T\mathbf{X}_i(s)} + e^{\hat{\xi}(s)}e^{\hat{\phi}^T\mathbf{Z}_i(s)}\right]}\left(\sum_{i=1}^{n} dN_{\bullet}(s)\right)$$

which is exactly the same as before, apart from a constant $e^{-\hat{\xi}(s)}$ in each contribution. This makes sence, since we use all observations to estimate $\Lambda_0^1(t)$, and we assume $h_1(t)e^{-\xi} = h_0(t)$.

Finally, an estimator for the all-cause cumulative baseline hazard is obtained by summing both previous estimators .

5 Selected software

Our method was implemented in the free software environment R (http://www.r-project.org/). The code consists of the implementation of the log likelihoods l and l^*, the various analytical expressions of the gradient components of these likelihoods (which may be used to obtain a faster optimization), the hessian and the $\mathbf{\Delta}$ derived in the previous section (which are used in obtaining the theoretical standard errors on estimates). We also provide an example of the simulation of data.

5.1 Required format of the data set

The data set is primarily built up as a matrix, with individuals arranged in rows and covariates in columns. The first columns of the matrix contain the covariates $\mathbf{Z}$, next the covariates $\mathbf{X}$, and then the observation times, overall status (0=censoring, 1=event), an observed-event-type indicator (0=censoring, 1=event-of-interest observation, 2=competing-risk observation), a type-1 indicator (1= type-1 observation, 0=else) and a type-0 indicator (1=competing-risk observation, 0=else). The last two columns contain the information regarding the prespecified misclassification rates. The format of this is a column `peen` which is equal to $1 - p_1$ for type-1 observations and p_1 for competing risks. Likewise, a column `pnul` is $1-p_0$ for type-0 observations and p_0 for events of interest. Both columns are set to 0 for censorings. By storing this data matrix as a data frame, the covariate names can be attached.

```
data=
 as.data.frame(cbind(covars,times,allstatus,obsstatus,obsstatus1,obsstatus0,peen,pnul))
names(data)=
 c(covarnames,"times","allstatus","obsstatus","obsstatus1","obsstatus0","peen","pnul")
```

The observations should be ordered by increasing observation time.

The later likelihood implementations expect this data set to be available both as a data frame `data` and as a matrix `datamat`.

5.2 Simulation of data

To clarify the required structure of the data, we provide the code for a small data simulation. This code assumes constant cause-specific hazards, which are both proportionally affected by the same binary treatment. A group of 200 patients is uniformly accrued into a trial over a period of 2 years, with an additional 4 year-period of follow-up. Only administrative censoring occurs, which is independent of survival here. Cause-of-death is assumed to be misclassified in 20% of the cases, for both causes.

The code includes small functions for ordering the data according to observation time and for adding the misclassifications probabilities to the data set. The function for ordering assumes no ties occur (and adds a minute quantity to censored data, to provide proper risk set definitions). The function for adding the misclassification probabilities works with constant misclassification rates independent of any covariates other than treatment.

Finally, for information on more complex simulation (e.g. time-varying hazards) see Beyersmann *and others*.[9]

```
# PROGRAM TO ORDER DATA BY FAILURE TIME #
 timeorder=function(data){
  data$times=data$times+0.000000000000001*(1-data$allstatus)   #at equal times censorings are put last
  return(data[sort.list(data$times,na.last=NA,method="quick"),])
 }
```

```
# PROGRAM TO ADD P_0 AND P_1 TO THE DATA SET #
# if wanted, these may depend on t, but not implemented here!
 padd=function(data){    #remember: allstatus is failure indicator, only 0 for censorings
  data$peen=data$allstatus*((1-p1)*data$obsstatus1+p1*(1-data$obsstatus1))
  data$pnul=data$allstatus*(p0*data$obsstatus1+(1-p0)*(1-data$obsstatus1))
  return(data)
 }

#DATA GENERATION
 n=200 # number of patients
 maxfollowup=6 # length of trial
 lengthaccrual=2 # length of accrual period
 nZ=1 # number of covariates acting on event of interest
 nX=1 # number of covariates acting on competing risk

# aim for overall hazard of 0.6 under control
# h0 and h1 are chosen to yield close to 50\% events under control
# h0+h1=h0(1+exp(xi))=0.6
# set relative cause-specific hazard and from that cause-specific hazards
 xi=-1
 h0=0.6/(1+exp(xi))
 h1=0.6-h0

# set covariate effects
 phi=1
 rho=-0.5

# set misclassification probabilities
 p0=0.2
 p1=0.2

# generate treatment covariate
 trt=rbinom(n,1,0.5)

# generate three survival times : type 0 = competing risk, type 1 = event of interest and censoring = type 2
 times0=rweibull(n,1,1/(h0*exp(trt*rho)))
 times1=rweibull(n,1,1/(h1*exp(trt*phi)))
 timesc=runif(n,maxfollowup-lengthaccrual,maxfollowup)

# create the minimum, the real failure type and the observed failure type
 times=pmin(times0,times1,timesc)                     #observe minimum time
 realstatus=(times==times1)+(times==timesc)*2         #REAL failure type as 0, 1 or 2
 realstatus1=(realstatus==1)                          #is it an event of interest? (0 or 1)
 realstatus0=(realstatus==0)                          #is it a competing risk? (0 or 1)
 changeprob=realstatus1*p1+realstatus0*p0             #define probability of misclassification
 change=rbinom(n,1,changeprob)                        #indicator of misclassification
 obsstatus=realstatus+change*(realstatus0-realstatus1)  #OBSERVED failure type as 0, 1 or 2
 obsstatus1=(obsstatus==1)                            #do we observe an event of interest? (0 or 1)
 obsstatus0=(obsstatus==0)                            #do we observe a competing risk? (0 or 1)
 allstatus=(realstatus==0 | realstatus==1)            #failure indicator (1=failure, 0=censoring)

# making the output data set
 data2=as.data.frame(cbind(trt,trt,times,allstatus,obsstatus,obsstatus1,obsstatus0))
 names(data2)=c("Z1","X1","times","allstatus","obsstatus","obsstatus1","obsstatus0")

# order data by observation time
 data=timeorder(data2)

# add misclassification probabilities to data set
# may be individual specific, but not implemented here
 data=padd(data)
 datamat=as.matrix(data)
```

5.3 Definition of the likelihoods

Remark: while the method supports case-specific misclassification rates, and these can be used in the definition of the data set, our implementations of the likelihoods only support misclassification rates that depend on event type and time alone!

The implementations of both l and l^* assume that a vector `input` exists, which contains the estimates of (in order) (ξ, ρ, ϕ) (or starting values for their optimization).

The function `loglik` implements minus the log likelihood l.

```
loglikopt=function(par){ ###par contains (rho,phi)
 xi=input[1]
 if(nX==0){rho=as.matrix(0)}else{rho=as.matrix(c(rep(0,nZ),par[1:nX]))}
 if(nZ==0){phi=as.matrix(0)}else{phi=as.matrix(c(par[(nX+1):(nZ+nX)],rep(0,nX)))}
 ll=-(sum(log((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)[data$pnul!=0]
       +(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen)[data$peen!=0]))
       +sum(-log((cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)[n:1]))[n:1]*data$pnul
       +cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)[n:1]))[n:1]*data$peen)[data$peen!=0])))
 ll
}
```

The function `logliksster` implements minus the log likelihood l^*.

```
logliksteropt=function(par){ ###par contains (xi)
 xi=par
 if(nX==0){rho=as.matrix(0)}else{rho=as.matrix(c(rep(0,nZ),input[2:(1+nX)]))}
 if(nZ==0){phi=as.matrix(0)}else{phi=as.matrix(c(input[(nX+2):(nZ+nX+1)],rep(0,nX)))}
 ll=-(sum(log((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)[data$pnul!=0]
       +(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen)[data$peen!=0]))
       +sum(-log((cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)[n:1]))[n:1]
       +cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)[n:1]))[n:1])[data$peen!=0])))
 ll
}
```

Finally, we also provide a function which can be used to find starting values for all parameters for the optimization.

```
loglik=function(par){ ###par contains (xi,rho,phi)
 xi=par[1]
 if(nX==0){rho=as.matrix(0)}else{rho=as.matrix(c(rep(0,nZ),par[2:(1+nX)]))}
 if(nZ==0){phi=as.matrix(0)}else{phi=as.matrix(c(par[(nX+2):(nZ+nX+1)],rep(0,nX)))}
 ll=-(sum(log((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)[data$pnul!=0]
  +(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen)[data$peen!=0]))
  +sum(-log((cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)[n:1]))[n:1]*data$pnul
  +cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)[n:1]))[n:1]*data$peen)[data$peen!=0])))
 ll
}
```

5.4 Definition of the gradients

While the gradient components are not needed for the optimization procedure as we use it, they may be used to enhance the efficiency of the optimization procedure.

```
grad=function(par){ ###par contains (xi,rho,phi)
 xi=par[1]
 if(nX==0){rho=as.matrix(0)}else{rho=as.matrix(c(rep(0,nZ),par[2:(1+nX)]))}
 if(nZ==0){phi=as.matrix(0)}else{phi=as.matrix(c(par[(nX+2):(nZ+nX+1)],rep(0,nX)))}
 grad=rep(NA,nZ+nX+1)
 weg=(data$peen!=0)
 # gradient of l*
 grad[1]=sum((((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)[weg])/((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)[weg]
  +(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen)[weg]))-(((cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho))[n:1])[n:1])[weg])
  /(cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)[n:1]))[n:1]+cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)[n:1]))[n:1])[weg]))
 # the gradient of l
 if(nZ!=0){
  for(nn in (nX+2):(nZ+nX+1)){
   dnn=nn-nX-1
   grad[nn]=-(sum(((datamat[,dnn]*(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen))[weg])
    /((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)[weg]+(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen)[weg]))
    -sum(((cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)*datamat[,dnn])[n:1])[n:1]*data$peen)[weg])
    /(cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)[n:1]))[n:1]*data$pnul
    +cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)[n:1]))[n:1]*data$peen)[weg]))
   }
  }
```

```
 if(nX!=0){
  for(mm in 2:(nX+1)){
   dmm=mm+nZ-1
   grad[mm]=-(sum(((datamat[,dmm]*(exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul))[weg])
   /((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)[weg]+(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen)[weg]))
   -sum(((cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*datamat[,dmm])[n:1])[n:1]*data$pnul)[weg])
   /((cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)[n:1]))[n:1]*data$pnul
   +cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)[n:1]))[n:1]*data$peen)[weg])))
  }
 }
return(grad)
}
```

5.5 Definition of the covariance matrix

First we define the hessian Γ.

```
hess=function(par){ ###par contains (xi,rho,phi)
 xi=par[1]
 if(nX==0){rho=as.matrix(0)}else{rho=as.matrix(c(rep(0,nZ),par[2:(1+nX)]))}
 if(nZ==0){phi=as.matrix(0)}else{phi=as.matrix(c(par[(nX+2):(nZ+nX+1)],rep(0,nX)))}
 Hessian=matrix(NA,nrow=nZ+nX+1,ncol=nZ+nX+1)
 weg=(data$peen!=0)
 if(nZ!=0){
 for(mm in (nX+2):(nZ+nX+1)){
  for(nn in mm:(nZ+nX+1)){ # this is component d^2l/dphi^2
   dmm=mm-nX-1
   dnn=nn-nX-1
   Hessian[mm,nn]=sum((((datamat[,dmm]*datamat[,dnn]*(exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)
    *(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen))[weg])/((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)[weg]
    +(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen)[weg])^2)-(((cumsum(((exp(datamat[,1:(nZ+nX)]%*%phi)*datamat[,dmm]
    *datamat[,dnn])[n:1]))[n:1]*data$peen)[weg])/(cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)[n:1]))[n:1]*data$pnul
    +cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)[n:1]))[n:1]*data$peen)[weg])+((((cumsum(((exp(datamat[,1:(nZ+nX)]%*%phi)
    *datamat[,dmm])[n:1]))[n:1]*data$peen)[weg])*((cumsum(((exp(datamat[,1:(nZ+nX)]%*%phi)*datamat[,dnn])[n:1]))[n:1]
    *data$peen)[weg]))/((cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)[n:1]))[n:1]*data$pnul
    +cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)[n:1]))[n:1]*data$peen)[weg])^2))
   Hessian[nn,mm]=Hessian[mm,nn] #symmetry within the dphi^2 block!
   }
  }
 }
 if(nX!=0){
 for(mm in 2:(nX+1)){
  for(nn in mm:(nX+1)){ # this is component d^2l/drho^2
   dmm=mm+nZ-1
   dnn=nn+nZ-1
   Hessian[mm,nn]=sum((((datamat[,dmm]*datamat[,dnn]*(exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)
   *(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen))[weg])/((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)[weg]
   +(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen)[weg])^2)-(((cumsum(((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*datamat[,dmm]
   *datamat[,dnn])[n:1]))[n:1]*data$pnul)[weg])/(cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)[n:1]))[n:1]*data$pnul
   +cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)[n:1]))[n:1]*data$peen)[weg])+((((cumsum(((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)
   *datamat[,dmm])[n:1]))[n:1]*data$pnul)[weg])*((cumsum(((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*datamat[,dnn])[n:1]))[n:1]
   *data$pnul)[weg]))/((cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)[n:1]))[n:1]*data$pnul
   +cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)[n:1]))[n:1]*data$peen)[weg])^2))
   Hessian[nn,mm]=Hessian[mm,nn] #symmetry within the drho^2 block!
   }
  }
 }
 # code for d^2l*/dxi^2
 Hessian[1,1]=sum(((((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)*(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen))[weg])
  /((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)[weg]+(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen)[weg])^2)
  -((((cumsum(((exp(-xi+datamat[,1:(nZ+nX)]%*%rho))[n:1]))[n:1])[weg])
  *((cumsum(((exp(datamat[,1:(nZ+nX)]%*%phi))[n:1]))[n:1])[weg]))/((cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)[n:1]))[n:1]
  +cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)[n:1]))[n:1])[weg])^2))
 if(nX!=0 & nZ!=0){
 for(mm in 2:(nX+1)){
  for(nn in (nX+2):(nZ+nX+1)){ # this is component d^2l/drhodphi
   dmm=mm+nZ-1
   dnn=nn-nX-1
   Hessian[mm,nn]=sum((((-datamat[,dmm]*datamat[,dnn]*(exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)
```

```
    *(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen))[weg])/((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)[weg]
    +(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen)[weg])^2)
    +((((cumsum(((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*datamat[,dmm])[n:1]))[n:1]*data$pnul)[weg])
    *((cumsum(((exp(datamat[,1:(nZ+nX)]%*%phi)*datamat[,dnn])[n:1]))[n:1]*data$peen)[weg]))
    /((cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)[n:1]))[n:1]*data$pnul
    +cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)[n:1]))[n:1]*data$peen)[weg])^2))
   Hessian[nn,mm]=Hessian[mm,nn] # symmetry within the drhodphi block!
   }
  }
 }
 mm=1
 if(nZ!=0){
  for(nn in (nX+2):(nZ+nX+1)){
   # first component d^2l*/dphidxi
   dnn=nn-nX-1
   Hessian[nn,mm]=sum((((datamat[,dnn]*(exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)
    *(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen))[weg])/((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)[weg]
    +(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen)[weg])^2)-((((cumsum(((exp(-xi+datamat[,1:(nZ+nX)]%*%rho))[n:1]))[n:1])[weg])
    *((cumsum(((exp(datamat[,1:(nZ+nX)]%*%phi)*datamat[,dnn])[n:1]))[n:1])[weg]))
    /((cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)[n:1]))[n:1]+cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)[n:1]))[n:1])[weg])^2))
   # not symmetric anymore!! next do component d^2l/dxidphi
   Hessian[mm,nn]=sum((((datamat[,dnn]*(exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)
    *(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen))[weg])/((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)[weg]
    +(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen)[weg])^2)
    -((((cumsum(((exp(-xi+datamat[,1:(nZ+nX)]%*%rho))[n:1]))[n:1]*data$pnul)[weg])
    *((cumsum(((exp(datamat[,1:(nZ+nX)]%*%phi)*datamat[,dnn])[n:1]))[n:1]*data$peen)[weg]))
    /((cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)[n:1]))[n:1]*data$pnul
    +cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)[n:1]))[n:1]*data$peen)[weg])^2))
   }
  }
 mm=1
 if(nX!=0){
  for(nn in 2:(1+nX)){ # this is component d^2l*/drhodxi
   dnn=nn+nZ-1
   Hessian[nn,mm]=sum((((-datamat[,dnn]*(exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)
    *(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen))[weg])/((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)[weg]
    +(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen)[weg])^2)
    +((((cumsum(((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*datamat[,dnn])[n:1]))[n:1])[weg])
    *((cumsum(((exp(datamat[,1:(nZ+nX)]%*%phi))[n:1]))[n:1])[weg]))/((cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)[n:1]))[n:1]
    +cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)[n:1]))[n:1])[weg])^2))
   # not symmetric anymore!! next do component d^2l/dxidrho
   Hessian[mm,nn]=sum((((-datamat[,dnn]*(exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)
    *(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen))[weg])/((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)[weg]
    +(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen)[weg])^2)
    +((((cumsum(((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*datamat[,dnn])[n:1]))[n:1]*data$pnul)[weg])
    *((cumsum(((exp(datamat[,1:(nZ+nX)]%*%phi))[n:1]))[n:1]*data$peen)[weg]))
    /((cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)[n:1]))[n:1]*data$pnul
    +cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)[n:1]))[n:1]*data$peen)[weg])^2))
  }
 }
return(Hessian)
}
```

Next, we define the matrix $\boldsymbol{\Delta}$.

```
delta=function(par){ ###par contains (xi,rho,phi)
xi=par[1]
if(nX==0){rho=as.matrix(0)}else{rho=as.matrix(c(rep(0,nZ),par[2:(1+nX)]))}
if(nZ==0){phi=as.matrix(0)}else{phi=as.matrix(c(par[(nX+2):(nZ+nX+1)],rep(0,nX)))}
Delta=matrix(NA,nrow=nZ+nX+1,ncol=nZ+nX+1)
 weg=(data$peen!=0)
 if(nZ!=0){
 for(mm in (nX+2):(nZ+nX+1)){
  for(nn in mm:(nZ+nX+1)){ # this is component <T[1],T[1]>
   dmm=mm-nX-1
   dnn=nn-nX-1
   Delta[mm,nn]=sum((((datamat[,dmm]*datamat[,dnn]*(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen)
    *(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen))[weg])/((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)[weg]
    +(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen)[weg])^2)
    -((((cumsum(((exp(datamat[,1:(nZ+nX)]%*%phi)*datamat[,dmm])[n:1]))[n:1]*data$peen)[weg])
    *((cumsum(((exp(datamat[,1:(nZ+nX)]%*%phi)*datamat[,dnn])[n:1]))[n:1]*data$peen)[weg]))
```

```
   /((cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)[n:1]))[n:1]*data$pnul
   +cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)[n:1]))[n:1]*data$peen)[weg])^2))
  Delta[nn,mm]=Delta[mm,nn] # symmetry!
  }
 }
}
if(nX!=0){
for(mm in 2:(nX+1)){
 for(nn in mm:(nX+1)){ # this is component <T[2],T[2]>
  dmm=mm+nZ-1
  dnn=nn+nZ-1
  Delta[mm,nn]=sum((((datamat[,dmm]*datamat[,dnn]*(exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)
   *(exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul))[weg])/((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)[weg]
   +(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen)[weg])^2)
   -((((cumsum(((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*datamat[,dmm])[n:1]))[n:1]*data$pnul)[weg])
   *((cumsum(((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*datamat[,dnn])[n:1]))[n:1]*data$pnul)[weg]))
   /((cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)[n:1]))[n:1]*data$pnul
   +cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)[n:1]))[n:1]*data$peen)[weg])^2))
  Delta[nn,mm]=Delta[mm,nn] # it's a symmetric matrix!
  }
 }
}
# code for <T[3],T[3]>
Delta[1,1]=sum(((((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)*(exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul))[weg])
 /((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)[weg]+(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen)[weg])^2)
 -((((cumsum(((exp(-xi+datamat[,1:(nZ+nX)]%*%rho))[n:1]))[n:1])[weg])
 *((cumsum(((exp(-xi+datamat[,1:(nZ+nX)]%*%rho))[n:1]))[n:1])[weg]))
 /((cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)[n:1]))[n:1]+cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)[n:1]))[n:1])[weg])^2))
if(nX!=0 & nZ!=0){
 for(mm in 2:(nX+1)){
  for(nn in (nX+2):(nZ+nX+1)){ # this is component <T[2],T[1]> (completely equal to -d^2l/dxidrho!)
   dmm=mm+nZ-1
   dnn=nn-nX-1
  Delta[mm,nn]=sum((((datamat[,dmm]*datamat[,dnn]*(exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)
   *(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen))[weg])/((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)[weg]
   +(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen)[weg])^2)
   -((((cumsum(((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*datamat[,dmm])[n:1]))[n:1]*data$pnul)[weg])
   *((cumsum(((exp(datamat[,1:(nZ+nX)]%*%phi)*datamat[,dnn])[n:1]))[n:1]*data$peen)[weg]))
   /((cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)[n:1]))[n:1]*data$pnul
   +cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)[n:1]))[n:1]*data$peen)[weg])^2))
  Delta[nn,mm]=Delta[mm,nn] # it's a symmetric matrix!
  }
 }
}
mm=1
if(nZ!=0){
 for(nn in (nX+2):(nZ+nX+1)){ # this is component <T[3],T[1]> (completely equal to -d^2l/dxidphi!)
  dnn=nn-nX-1
  Delta[mm,nn]=sum((((-datamat[,dnn]*(exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)
   *(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen))[weg])/((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)[weg]
   +(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen)[weg])^2)
   +((((cumsum(((exp(-xi+datamat[,1:(nZ+nX)]%*%rho))[n:1]))[n:1]*data$pnul)[weg])
   *((cumsum(((exp(datamat[,1:(nZ+nX)]%*%phi)*datamat[,dnn])[n:1]))[n:1]*data$peen)[weg]))
   /((cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)[n:1]))[n:1]*data$pnul
   +cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)[n:1]))[n:1]*data$peen)[weg])^2))
  Delta[nn,mm]=Delta[mm,nn] # it's a symmetric matrix!
  }
 }
mm=1
if(nX!=0){
 for(nn in 2:(1+nX)){ # this is component <T[3],T[2]>
  dnn=nn+nZ-1
  Delta[mm,nn]=sum((((-datamat[,dnn]*(exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)
   *(exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul))[weg])/((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)[weg]
   +(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen)[weg])^2)
   +((((cumsum(((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)*datamat[,dnn])[n:1]))[n:1]*data$pnul)[weg])
   *((cumsum(((exp(-xi+datamat[,1:(nZ+nX)]%*%rho))[n:1]))[n:1]*data$pnul)[weg]))
   /((cumsum((exp(-xi+datamat[,1:(nZ+nX)]%*%rho)[n:1]))[n:1]*data$pnul
   +cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)[n:1]))[n:1]*data$peen)[weg])^2))
  Delta[nn,mm]=Delta[mm,nn] # it's a symmetric matrix!
 }
```

```
 }
return(Delta)
}
```

5.6 Example code for analysis

We next provide an example of an analysis.

```
n=dim(datamat)[1]
input=optim(par=c(2,-0.13,-0.13),loglik)$par
   # starting values which are chosen based on naive cox models and common sense
   # input contains (xi,rho,phi), in that order
input[1]=optim(par=input[1],logliksteropt,method = "L-BFGS-B",lower=rep(-10, 7), upper=rep(10, 7))$par
input[2:(nZ+nX+1)]=optim(par=input[2:(nZ+nX+1)],loglikopt,method = "CG")$par
input[1]=optim(par=input[1],logliksteropt,method = "L-BFGS-B",lower=rep(-10, 7), upper=rep(10, 7))$par
input[2:(nZ+nX+1)]=optim(par=input[2:(nZ+nX+1)],loglikopt,method = "CG")$par
input[1]=optim(par=input[1],logliksteropt,method = "L-BFGS-B",lower=rep(-10, 7), upper=rep(10, 7))$par
input[2:(nZ+nX+1)]=optim(par=input[2:(nZ+nX+1)],loglikopt,method = "CG")$par
input[1]=optim(par=input[1],logliksteropt,method = "L-BFGS-B",lower=rep(-10, 7), upper=rep(10, 7))$par
input[2:(nZ+nX+1)]=optim(par=input[2:(nZ+nX+1)],loglikopt,method = "CG")$par

h=hess(input)
d=delta(input)
diag(solve(h)%*%d%*%t(solve(h))) # this yields the variance-covariance matrix
2*pnorm(-abs(input)/sqrt(diag(solve(h)%*%d%*%t(solve(h))))) # this yields a Wald-test
```

5.7 Time-varying analysis

The time varying analysis revolves around the following expression:

```
smoothxicovars=function(expminxi){
 if(nX==0){rho=as.matrix(0)}else{rho=as.matrix(c(rep(0,nZ),input[2:(1+nX)]))}
 if(nZ==0){phi=as.matrix(0)}else{phi=as.matrix(c(input[(nX+2):(nZ+nX+1)],rep(0,nX)))}
 -sum(exp(-(time-data$times)^2/(2*h*h))*(data$allstatus*log(1/(expminxi*exp(datamat[,1:(nZ+nX)]%*%rho)
 +exp(datamat[,1:(nZ+nX)]%*%phi)))+data$obsstatus1*log(exp(datamat[,1:(nZ+nX)]%*%phi)*(1-p1)
 +expminxi*exp(datamat[,1:(nZ+nX)]%*%rho)*p0)+data$obsstatus0*log((1-p0)*exp(datamat[,1:(nZ+nX)]%*%rho)*expminxi
 +p1*exp(datamat[,1:(nZ+nX)]%*%phi))))
 }
```

The optimization of this is iterated until convergence with the optimization of the following expression (which is a slightly modified version of the earlier function `loglikopt`, to allow for a time-varying ξ):

```
loglikoptxi=function(par){ ###par contains (rho,phi)
 if(nX==0){rho=as.matrix(0)}else{rho=as.matrix(c(rep(0,nZ),par[1:nX]))}
 if(nZ==0){phi=as.matrix(0)}else{phi=as.matrix(c(par[(nX+1):(nZ+nX)],rep(0,nX)))}
 ll=-(sum(log((exp(log(xis)+datamat[,1:(nZ+nX)]%*%rho)*data$pnul)[data$pnul!=0]
  +(exp(datamat[,1:(nZ+nX)]%*%phi)*data$peen)[data$peen!=0]))
  +sum(-log((xis*+cumsum((exp(datamat[,1:(nZ+nX)]%*%rho)[n:1]))[n:1]*data$pnul
  +cumsum((exp(datamat[,1:(nZ+nX)]%*%phi)[n:1]))[n:1]*data$peen)[data$peen!=0])))
 ll
}
```

References

[1] Schoenfeld D. The Asymptotic Properties of Nonparametric Tests for Comparing Survival Distributions. *Biometrika.* 1981; **68**:316–319.

[2] Pintilie M. *Competing Risks : A Practical Perspective.* Chichester: John Wiley & Sons, Inc.; 2006.

[3] Aalen OO, Borgan 0, Gjessing HK. *Survival and Event History Analysis*. New York: Springer; 2008.

[4] Andersen PK, Gill R. Cox's regression model for counting processes: a large sample study (with commentary). *Annals of Statistics*. 1982; **10**:1100–1124.

[5] Goetghebeur E, Ryan L. Analysis of Competing Risks Survival Data When Some Failure Types are Missing. *Biometrika*. 1995; **82**:821–833.

[6] Van Rompaye B, Jaffar S, Goetghebeur E. Design and testing for clinical trials faced with misclassified causes of death. *Biostatistics*. 2010 **11**:546–558.

[7] Fleming TR, Harrington DP. *Counting processes and survival analysis*. John Wiley & Sons, Inc., New York; 1991.

[8] Dewanji A. A Note on a Test for Competing Risks with Missing Failure Type. *Biometrika*. 1992; **79**:855–857.

[9] Beyersmann J, Latouche A, Buchholz A, Schumacher M. Simulating competing risks data in survival analysis. *Statistics in Medicine*. 2009 **28**:956–971.