

**Web Appendix for “Restricted mean survival time to estimate  
an intervention effect in a cluster randomized trial”**

*Le Vilain--Abraham F, Tavernier E, Dantan E, Desmée S, Caille A*

## Web Appendix A: Generalized estimating equation

### A.1. Generalized estimating equation and sandwich variance estimator for the pseudo-values regression for independent data

In section 2.2.3 in the main text, we consider the following generalized linear regression model:

$$g(E[\min(T_l, t^*) | \tilde{\mathbf{Z}}_l]) = \boldsymbol{\beta}' \tilde{\mathbf{Z}}_l \quad (1)$$

where  $t^*$  is the time horizon,  $T_l$  is the true time-to-event for individual  $l$ ;  $\tilde{\mathbf{Z}}_l = (1, Z_l)'$ , where  $Z_l$  is the intervention variable = 1 if the individual  $l$  received the intervention and 0 if the individual is in a control group;  $\boldsymbol{\beta} = (\beta_0, \beta_1)'$ , where  $\beta_0$  is the intercept and  $\beta_1$  is the intervention effect, and  $g(\cdot)$  a link function. Because the survival time can be censored,  $\min(T_l, t^*)$  is not observed for all individuals and is replaced by the pseudo-values  $Y_l$ , defined in equation (6) in the main text, in equation (1). The regression coefficients may be estimated from the following generalized estimating equation (GEE):<sup>1,2</sup>

$$\mathbf{U}(\boldsymbol{\beta}) = \sum_{l=1}^N \mathbf{U}_l(\boldsymbol{\beta}) = \sum_{l=1}^N \frac{\partial \mu_l}{\partial \boldsymbol{\beta}} V_l^{-1} (\hat{Y}_l - \mu_l) = 0 \quad (2)$$

where  $V_l$  is a working variance and  $\mu_l = g^{-1}(\boldsymbol{\beta}' \mathbf{Z}_l)$ . Andersen et al.<sup>3</sup> showed that the estimates of the regression coefficients are consistent and the variance for  $\hat{\boldsymbol{\beta}}$  could be estimated with the following sandwich estimator:

$$\boldsymbol{\Sigma} = \mathbf{I}(\hat{\boldsymbol{\beta}})^{-1} \widehat{\text{var}}(\mathbf{U}(\hat{\boldsymbol{\beta}})) \mathbf{I}(\hat{\boldsymbol{\beta}})^{-1} \quad (3)$$

where

$$\mathbf{I}(\boldsymbol{\beta}) = \sum_{l=1}^N \left( \frac{\partial \mu_l}{\partial \boldsymbol{\beta}} \right)' V_l^{-1} \left( \frac{\partial \mu_l}{\partial \boldsymbol{\beta}} \right),$$

$$\widehat{\text{var}}(\mathbf{U}(\hat{\boldsymbol{\beta}})) = \sum_{l=1}^N \mathbf{U}_l(\hat{\boldsymbol{\beta}}) \mathbf{U}_l(\hat{\boldsymbol{\beta}})'$$

A.2. *Generalized estimating equation and exchangeable working correlation matrix: estimation of  $\rho$  and  $\phi$  parameters*

In section 2.3.3 in the main text, we consider a sample of  $K$  clusters of size  $m_k$  ( $k=1, \dots, K$ ) and then  $n = \sum_{k=1}^K m_k$  individuals. The following generalized estimating equation is considered:

$$\mathbf{U}(\boldsymbol{\beta}) = \sum_{k=1}^K \mathbf{U}_k(\boldsymbol{\beta}) = \sum_{k=1}^K \left( \frac{\partial \boldsymbol{\mu}_k}{\partial \boldsymbol{\beta}} \right)' \mathbf{V}_k^{-1} (\hat{\mathbf{Y}}_k - \boldsymbol{\mu}_k) = 0 \quad (4)$$

where  $\hat{\mathbf{Y}}_k = (\hat{Y}_{k1}, \dots, \hat{Y}_{km_k})$ ,  $\mu_{kl} = g^{-1}(\boldsymbol{\beta}' \tilde{\mathbf{Z}}_{kl})$  with  $\boldsymbol{\mu}_k = (\mu_{k1}, \dots, \mu_{km_k})$  and  $\mathbf{V}_k$  is an  $m_k \times m_k$  working covariance matrix for cluster  $k$ . It is defined as  $\mathbf{V}_k = \phi \mathbf{A}_k^2 \mathbf{R}(\rho) \mathbf{A}_k^2$ , where  $\phi$  is a dispersion parameter,  $\mathbf{R}(\rho)$  is an  $m_k \times m_k$  working correlation matrix and  $\mathbf{A}_k$  an  $m_k \times m_k$  diagonal matrix with  $\text{Var}(\mu_{kl})$  as diagonal elements. If the working

correlation matrix is exchangeable, then  $\mathbf{R}(\rho) = \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho \\ \rho & \cdots & \rho & 1 \end{pmatrix}$  where,  $\rho =$

$\text{Corr}(\hat{Y}_{kl}, \hat{Y}_{ki}) \forall k, i \neq l$ . The parameters  $\rho$  and  $\phi$  are estimated from the Pearson residuals defined as following:

$$\hat{r}_{kl} = \frac{\hat{Y}_{kl} - \mu_{kl}}{\sqrt{\text{Var}(\mu_{kl})}} \quad (5)$$

The dispersion parameter  $\phi$  is estimated by

$$\hat{\phi} = \frac{1}{n - p} \sum_{k=1}^K \sum_{l=1}^{m_k} \hat{r}_{kl}^2 \quad (6)$$

The parameter  $\rho$  is estimated by

$$\hat{\rho} = \frac{1}{(n^* - p)\phi} \sum_{k=1}^K \sum_{i \neq k} \hat{r}_{ki} \hat{r}_{ki} \quad (7)$$

where  $n^* = \sum_{k=1}^K m_k(m_k - 1)$ .

## Web Appendix B: Permutation-based confidence intervals

We consider the model (1) described in the Appendix A.1, with  $g(\cdot)$  the identity link function, for the pseudo-values regression.

$$E[\min(T_1, t^*) | \tilde{\mathbf{Z}}_1] = \boldsymbol{\beta}' \tilde{\mathbf{Z}}_1 = \beta_0 + \beta_1 Z_1 \quad (8)$$

A 100(1- $\alpha$ )% permutation-based confidence interval for the intervention effect  $\beta_1$  can be obtained by inverting the permutation test. This requires testing the non-zero null hypothesis  $H_0: \beta_1 = \beta^*$ , where  $\beta^* \neq 0$ , with permutation tests and collecting the set of values that are not rejected by these tests. To do so, Rabideau and Wang<sup>4</sup> proposed to rewrite the model (8) as follows:

$$E[\min(T_1, t^*) | \tilde{\mathbf{Z}}_1] = \beta_0 + \beta^* Z_1 + (\beta_1 - \beta^*) Z_1 = \beta_0 + \beta^* Z_1 + \delta_{\text{offset}} Z_1 \quad (9)$$

and to test an equivalent zero null hypothesis  $H_0: \delta_{\text{offset}} = (\beta_1 - \beta^*) = 0$ . The permutation test is carried out comparing the test statistic obtained by fitting model (9) using the observed dataset and the permuted test statistics obtained by fitting the following model under different permuted intervention allocations:

$$E[\min(T_1, t^*) | \tilde{\mathbf{Z}}_1] = \beta_0 + \beta^* Z_1^{\text{obs}} + \delta_{\text{offset}} Z_1^{\text{perm}} \quad (10)$$

with the fixed offset term  $\beta^* Z_1^{\text{obs}}$  where  $Z_1^{\text{obs}}$  is the intervention allocation in the observed dataset, and  $Z_1^{\text{perm}}$  is a permuted intervention allocation.

This method could be computationally intensive because permutation tests are performed on a large number of  $\beta^*$  and each permutation test is carried out on many permuted intervention allocations. Rabideau and Wang<sup>4</sup> proposed an efficient procedure, based on a method suggested by Garthwaite<sup>5</sup>, to reduce the computation time. It is a sequential search, performed separately for the lower and upper bounds, which considers only a single permuted intervention allocation for each permutation test. We computed permutation-based confidence interval using this approach, which is briefly described below. For further details, see Rabideau and Wang<sup>4</sup> and Garthmaite<sup>5</sup>.

### *B.1. Search procedure*

Suppose we considered a I-step search for the upper bound of the  $100(1-\alpha)\%$  confidence interval for  $\beta_1$ , denoted U. At the step i, we set  $\beta^* = U_i$ , the current value of U, and randomly permuted the allocation of the clusters to the control and intervention groups. We fit the model (10) to obtain the permuted test statistic  $T_i^{\text{perm}}$ . We also fit the model (9) with the observed dataset to obtain the observed test statistic  $T_i^{\text{obs}}$ . The upper bound is updated based on the observed and permuted test statistics.

$$U_{i+1} = \begin{cases} U_i - \frac{c(\alpha/2)}{i}, & \text{if } T_i^{\text{perm}} > T_i^{\text{obs}} \\ U_i + \frac{c(1 - \alpha/2)}{i}, & \text{if } T_i^{\text{perm}} \leq T_i^{\text{obs}} \end{cases}$$

where c is a positive constant, called the step length constant.

Independently, the search for the lower bound L is carried out in a similar way, using

$$L_{i+1} = \begin{cases} L_i + \frac{c(\alpha/2)}{i}, & \text{if } T_i^{\text{perm}} < T_i^{\text{obs}} \\ L_i - \frac{c(1 - \alpha/2)}{i}, & \text{if } T_i^{\text{perm}} \geq T_i^{\text{obs}} \end{cases}$$

The final updated value, at the step I, are used as the bounds of the CI.

## *B.2. Implementation*

Following the recommendations of Rabideau and Wang<sup>4</sup> and Garthwaite<sup>5</sup>, at the step  $i$ , we set  $c = \kappa(U_i - \hat{\beta}_1)$  and  $c = \kappa(\hat{\beta}_1 - L_i)$  for the upper and lower bounds, respectively, where  $\kappa = 2/(z_{\alpha/2}(2\pi)^{-\frac{1}{2}}\exp(-z_{\alpha/2}^2)/2)$  with  $\alpha$  the significance level of the permutation test,  $z_{\alpha/2}$  the  $\alpha/2$ -quantile of the standard normal distribution and  $\hat{\beta}_1$  the estimation of  $\beta_1$  obtained by fitting the model (8) with the observed dataset. The lower and upper bounds were initialized at  $\hat{\beta}_1 - \frac{t_2 - t_1}{2}$  and  $\hat{\beta}_1 + \frac{t_2 - t_1}{2}$ , respectively, where  $t_1$  and  $t_2$  are the second smallest and second largest permuted estimations of  $\delta_{\text{offset}}$  from a permutation test testing  $H_0: \beta_1 = \hat{\beta}_1$ , with  $\lceil(4 - \alpha)/\alpha\rceil$  permuted allocations. To avoid an early large change,  $i$  is set to  $\min(\lceil 0.3(4 - \alpha)/\alpha \rceil, 50)$  at the first step. Finally, we set I to 5000 iterations to obtain an asymptotic variance of the coverage of the 95% confidence interval of  $1.3 \times 10^{-5}$ .

## **Web Appendix C: Additional details for the simulation study**

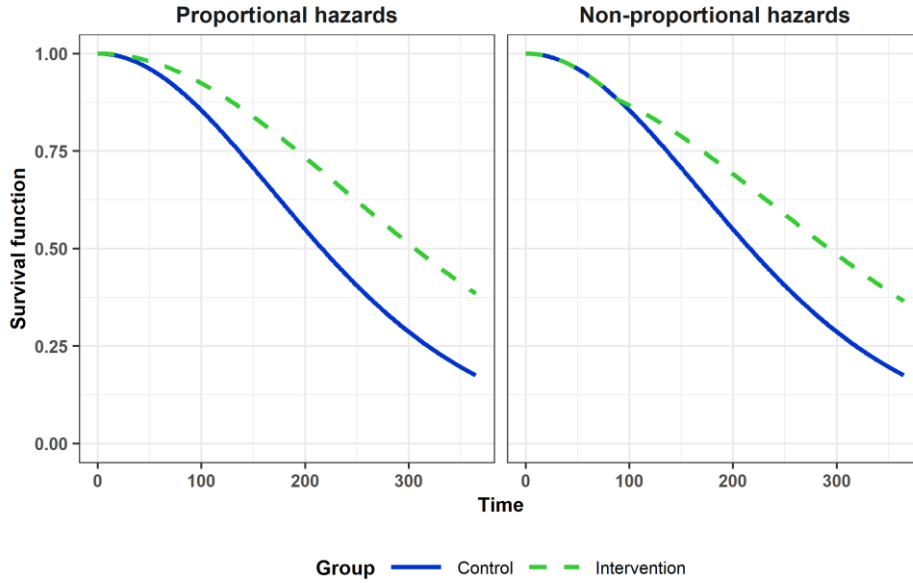
### *C.1. Simulation of censoring times*

We simulated random censoring times independent of the event times. We assumed that the censoring times in a cluster were independent. Considering a fixed censoring rate  $\psi$ , we generated  $R_{kl}$  (for  $l = 1, \dots, m_k, k = 1, \dots, K$ ) from a continuous uniform distribution on 0 to 1 and censored the observation if  $R_{kl} < \psi$ . When an observation

was censored, the censoring time for this individual  $l$  from the cluster  $k$  was generated with a uniform distribution on 0 to  $T_{kl}$ , the time-to-event of this individual.<sup>6</sup>

### C.2. Simulated survival function

Figure C2: Simulated survival functions under proportional and non-proportional hazards assumptions for true HR = 0.5 and Kendall's tau  $\tau = 0.1$ . The survival functions for the other scenarios are similar.



### C.3. Computation of the true difference in RMST

We computed the true difference in RMST up to  $t^*$  by integrating the marginal survival function in each group between 0 and  $t^*$ .

$$\int_0^{t^*} S(t, Z = 1) dt - \int_0^{t^*} S(t, Z = 0) dt \quad (8)$$

where  $S(t, Z = j)$  is the marginal survival function for the group  $j$  ( $j=0$  for the control group, 1 for the intervention group). The marginal survival function in each group was derived from the conditional survival function by integrating out the frailty with respect to the gamma density function.<sup>7</sup>

$$S(t, Z = j) = \int_0^{+\infty} S(t, Z = j | u_k) f(u_k) du_k$$

where  $S(t, Z = j | u_k)$  is the conditional survival function for individuals in group  $j$  ( $j = 0,1$ ) and  $f(\cdot)$  is the density function of the distribution gamma  $\Gamma\left(\frac{1}{\theta}, \frac{1}{\theta}\right)$  with a mean of 1 and variance  $\theta$ . An analytic form was obtained by using the Laplace transform.<sup>7</sup>

$$S(t, Z = j) = \mathcal{L}(H(t)) = \frac{1}{\left(1 + \frac{\lambda t^\rho \exp(Z\beta_G)}{1/\theta}\right)^{1/\theta}}$$

where  $\mathcal{L}(\cdot)$  is the Laplace transform and  $H(\cdot)$  the cumulative hazard function.

The integrals in (4) were approximated by using the numerical integration method of the right Riemann sum.

For non-proportional hazards assumption, we summed the true  $\Delta$ RMST computed separately over the 2 periods (0 to 90 days, 90 days to  $t^*$ ) with the corresponding treatment effect  $\beta$ .

Table C3: True difference in RMST up to 365 days for each scenario setting

Hazards assumption	True hazard ratio $\exp(\beta_G)$	Kendall's tau $\tau$	True $\Delta$ RMST up to 365
Null hypothesis	1	0.001	0
		0.01	0
		0.05	0
		0.1	0
		0.2	0
Proportional hazards	0.5	0.001	55.15
		0.01	54.78
		0.05	53.11
		0.1	51.00
		0.2	46.70
	0.8	0.001	18.72
		0.01	18.60
		0.05	18.04
		0.1	17.33



		0.2	15.87
Non-proportional hazards	0.5 (after 90 days)	0.001	42.03
		0.01	41.72
		0.05	40.33
		0.1	38.58
		0.2	35.01

#### C.4. Definition of the performance measures reported in the simulations

Let's denote  $\widehat{\Delta RMST}(t^*)_d$  and  $\widehat{V}(\widehat{\Delta RMST}(t^*)_d)$  the  $\widehat{\Delta RMST}(t^*)$  and its variance estimation in the simulated dataset  $d$  ( $d = 1, \dots, 1000$ ) and  $\Delta_{RMST}^{t^*}$  the true RMST difference. The performance of methods maintaining the type I error rate was assessed with the following criteria:

(1) The relative bias calculated as  $\frac{\overline{\Delta RMST}(t^*) - \Delta_{RMST}^{t^*}}{\Delta_{RMST}^{t^*}} \times 100$ , where  $\overline{\Delta RMST}(t^*) =$

$$\frac{1}{1000} \sum_{d=1}^{1000} \widehat{\Delta RMST}(t^*)_d,$$

(2) the relative error (RE) of the asymptotic standard errors (ASE), defined as

$$\frac{ASE - ESE}{ESE} \times 100, \text{ where } ASE = \sqrt{\frac{1}{1000} \sum_{d=1}^{1000} \widehat{V}(\widehat{\Delta RMST}(t^*)_d)}$$

$$\text{error } ESE = \sqrt{\frac{1}{999} \sum_{d=1}^{1000} (\widehat{\Delta RMST}(t^*)_d - \overline{\Delta RMST}(t^*))^2}.$$

The RE corresponds to the bias in the estimation of the true standard error, and thus the variance, of the  $\widehat{\Delta RMST}(t^*)$ . A value of the RE close to 0 reflects a good estimation of the asymptotic standard error of the intervention effect. When the RE deviates from 0, the variance is underestimated if RE is  $< 0$  and overestimated otherwise.

(3) The empirical coverage rate of the nominal 95% confidence intervals defined as the proportion of 95% confidence intervals containing the true RMST difference.

## References

1. Andersen PK, Hansen MG and Klein JP. Regression analysis of restricted mean survival time based on pseudo-observations. *Lifetime Data Anal* 2004; 10: 335- 50.
2. Andersen PK and Perme MP. Pseudo-observations in survival analysis. *Stat Methods Med Res* 2010; 19: 71- 99.
3. Andersen P, Klein J and Rosthøj S. Generalised linear models for correlated pseudo-observations, with applications to multi-state models. *Biometrika* 2003; 90: 15-27.
4. Rabideau DJ and Wang R. Randomization-based confidence intervals for cluster randomized trials. *Biostatistics* 2021; 22: 913-927.
5. Garthwaite, PH. Confidence Intervals from Randomization Tests. *Biometrics* 1996; 52: 1387–13936. Srivastava DK, George EO, Lu Z, et al. Impact of unequal censoring and insufficient follow-up on comparing survival outcomes: Applications to clinical studies. *Stat Methods Med Res* 2021; 30: 2057-2074.
7. Duchateau L and Janssen P. *The frailty model*. New York: Springer Verlag, 2008.

## Web Appendix D: Convergence

Table D1: Scenario settings and percentage of simulation iterations that did not converge for the pseudo-values regression with an exchangeable correlation matrix. There were no convergence issues for the scenarios not reported in the Table.

Total number of clusters	Mean cluster size	Kendall's tau $\tau$	True hazard ratio $exp(\beta_G)$	Percentage of datasets with convergence issues
<i>Null hypothesis</i>				
10	80	0.001	1	14.0
10	80	0.01	1	3.0
10	80	0.05	1	0.2
20	80	0.001	1	6.4
20	80	0.01	1	0.2
40	80	0.001	1	2.0
50	25	0.001	1	1.4
50	80	0.001	1	0.6
100	25	0.001	1	0.2
<i>Proportional hazards</i>				
10	80	0.001	0.8	14.9
10	80	0.001	0.5	15.4
10	80	0.01	0.8	3.2
10	80	0.01	0.5	5.4
10	80	0.05	0.8	0.2
10	80	0.05	0.5	0.2
10	80	0.1	0.5	0.1
20	80	0.001	0.8	4.2
20	80	0.001	0.5	4.4
20	80	0.01	0.8	0.4
20	80	0.01	0.5	0.7
40	80	0.001	0.8	1.5
40	80	0.001	0.5	1.6
50	25	0.001	0.8	1.6
50	25	0.001	0.5	1.0
50	25	0.01	0.8	0.2
50	80	0.001	0.8	0.3
50	80	0.001	0.5	0.6
100	25	0.001	0.8	0.2
100	25	0.001	0.5	0.2
<i>Non-proportional hazards</i>				
10	80	0.001	0.5 (after 90 days)	14.3
10	80	0.01	0.5 (after 90 days)	4.4
10	80	0.05	0.5 (after 90 days)	0.2
20	80	0.001	0.5 (after 90 days)	5.2
20	80	0.01	0.5 (after 90 days)	0.5
40	80	0.001	0.5 (after 90 days)	1.2
50	25	0.001	0.5 (after 90 days)	1.5
50	25	0.01	0.5 (after 90 days)	0.5
50	80	0.001	0.5 (after 90 days)	0.3
100	25	0.001	0.5 (after 90 days)	0.2
100	25	0.01	0.5 (after 90 days)	0.1
100	80	0.001	0.5 (after 90 days)	0.2

## Web Appendix E: Simulation results for the type I error rate

Table E1: Type I error rate (%) under the proportional hazards assumption for different combinations of number of clusters ( $K = 50, 100$ ) and Kendall's tau ( $\tau = 0.001, 0.01, 0.05, 0.1, 0.2$ ) with mean cluster size  $m = 25$ . Type I error between 3.6% and 6.4%, indicated in bold, was considered acceptable based on a binomial model with 1000 simulated datasets.

Number of clusters $K$	Kendall's $\tau$	Methods for independent data		Extended methods for clustered data		
		KM <sub>indep</sub>	PV <sub>indep</sub>	KM <sub>clust</sub>	PV <sub>ECM</sub>	PV <sub>ICM</sub>
50	0.001	<b>5.2</b>	<b>5.1</b>	<b>5.6</b>	6.9	<b>5.8</b>
	0.01	10.8	10.7	<b>6.2</b>	7.3	<b>6.3</b>
	0.05	25.3	25.3	<b>5.5</b>	<b>5.9</b>	<b>6.1</b>
	0.1	39.9	39.9	<b>5.4</b>	<b>6.0</b>	<b>5.8</b>
	0.2	51.6	51.6	7.4	7.8	7.7
100	0.001	<b>6.0</b>	<b>6.0</b>	<b>6.1</b>	<b>5.8</b>	<b>5.9</b>
	0.01	11.5	11.5	<b>5.3</b>	<b>5.2</b>	<b>5.6</b>
	0.05	27.3	27.3	<b>6.1</b>	<b>5.7</b>	<b>6.0</b>
	0.1	39.1	39.1	<b>6.1</b>	<b>5.6</b>	6.5
	0.2	50.1	50.1	6.7	<b>5.6</b>	7.1

## Web Appendix F: Simulation results with mean cluster size $m = 25$

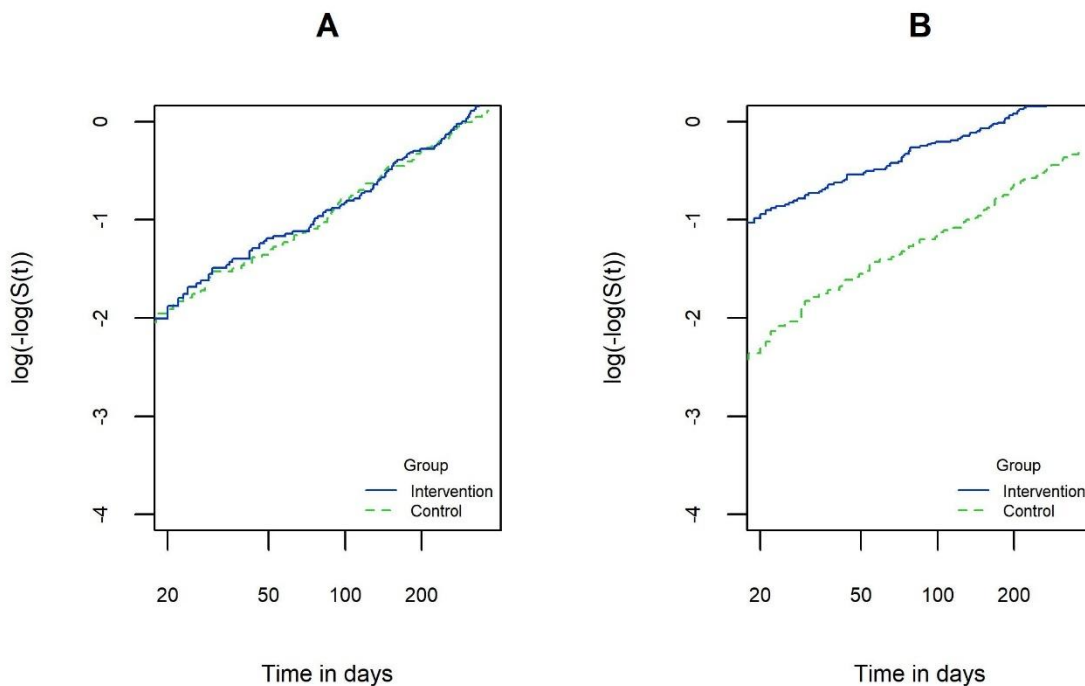
Table F1: Relative bias (%), relative error (%) and coverage rate (%) under proportional and non-proportional hazards assumptions for different combinations of number of clusters ( $K = 50, 100$ ), Kendall's tau ( $\tau = 0.001, 0.01, 0.05, 0.1, 0.2$ ) and true HR (HR = 0.8, 0.5) with mean cluster size  $m = 25$ .

True HR	K	$\tau$	Relative bias			Relative error			Coverage rate		
			KM <sub>clust</sub>	PV <sub>ECM</sub>	PV <sub>ICM</sub>	KM <sub>clust</sub>	PV <sub>ECM</sub>	PV <sub>ICM</sub>	KM <sub>clust</sub>	PV <sub>ECM</sub>	PV <sub>ICM</sub>
0.8	50	0.001	-2.29	-2.20	-2.31	0.4	-1.4	-0.4	93.6	92.4	93.1
		0.01	2.08	2.13	2.05	-1.5	-2.9	-2.1	94.3	93.9	94.1
		0.05	-0.19	-0.53	-0.23	-5.0	-3.2	-5.1	91.3	92.2	91.2
		0.1	0.18	0.12	0.15	-4.8	-1.0	-4.7	93.0	93.5	92.5
		0.2	-4.01	-0.83	-4.05	-6.4	-3.8	-6.3	93.1	92.7	91.8
	100	0.001	1.11	1.16	1.08	-2.4	-3.3	-2.7	94.0	93.3	94.1
		0.01	-0.70	-0.68	-0.72	-5.5	-6.0	-5.6	93.4	93.4	93.2
		0.05	-0.68	-0.75	-0.70	-0.2	1.5	-0.1	94.6	94.3	94.6
		0.1	-5.19	-4.36	-5.21	0.0	0.1	0.1	93.9	93.7	93.7
		0.2	-0.36	-0.55	-0.38	-2.3	-1.8	-2.1	94.7	94.7	94.5
0.5	50	0.001	-1.27	-1.51	-1.42	0.0	-3.2	-0.2	93.7	93.3	93.7
		0.01	-0.82	-0.98	-0.94	-6.4	-6.6	-6.7	93.0	93.0	93.0
		0.05	-1.94	-1.61	-2.05	-2.9	-2.0	-3.3	94.1	94.2	93.5
		0.1	-2.48	-2.61	-2.63	-6.3	-6.1	-6.4	92.6	92.9	91.9
		0.2	-2.78	-2.89	-2.90	-6.2	-6.3	-6.5	93.0	92.9	91.8
	100	0.001	-0.45	-0.57	-0.59	-1.7	-2.4	-1.4	94.0	94.0	94.1
		0.01	-0.56	-0.67	-0.69	2.6	2.7	2.8	95.7	95.2	95.6
		0.05	-1.04	-1.10	-1.15	0.2	0.3	0.1	94.9	94.7	95.2
		0.1	-2.14	-2.16	-2.24	-2.4	-4.0	-2.6	94.2	92.8	93.9
		0.2	-3.58	-3.62	-3.67	-3.2	-2.2	-3.4	93.1	93.9	93.0
<i>Non-proportional hazards</i>											
0.5 (after 90 days)	50	0.001	2.58	2.59	2.47	-2.8	-4.8	-3.0	92.9	91.6	92.7
		0.01	2.64	2.74	2.55	-5.4	-5.6	-5.6	92.8	92.4	92.9
		0.05	1.69	1.67	1.59	-1.9	-1.1	-1.9	93.8	94.0	93.5
		0.1	2.09	2.57	1.99	-3.1	-2.4	-3.2	93.0	92.4	92.7
		0.2	0.72	0.91	0.62	-4.8	-3.5	-4.8	93.1	93.1	92.1
	100	0.001	2.35	2.27	2.24	-1.0	-0.4	-0.8	93.4	93.9	93.6
		0.01	3.10	2.96	3.01	-1.4	-2.3	-1.2	93.2	92.5	93.1
		0.05	1.62	1.69	1.56	-2.2	-1.7	-2.0	94.3	93.9	94.2
		0.1	1.81	1.98	1.72	-1.7	-3.5	-1.5	94.6	94.0	94.4
		0.2	1.07	0.94	0.99	-1.9	1.1	-1.8	94.4	94.8	93.6

## Web Appendix G: Checking the proportional hazards assumption and R code for the illustrative example OEDIPUS

### G.1. Checking the proportional hazards assumption

Figure G1: The  $\log(-\log(S(t)))$  curves for the outcomes of the OEDIPUS trial: A) time to first unscheduled contact with an asthma exacerbation; B) time to first asthma review in primary care



### G.2. R code to estimate the difference in RMST

The R code used to compute the difference in RMST and its confidence interval for the 2 outcomes of the OEDIPUS trial with all the proposed methods in Table 5 is detailed in the following section. The “ci” function used to estimate the permutation-based confidence interval is available on <https://github.com/Le-Vilain-Abraham/Simulation-study-with-R-software/tree/main/Proportional-hazards/1%20-%20Generation%20of%20the%20datasets%20and%20statistical%20analysis/Permutation%20test/confidence%20interval>. The functions called by the “ci” function (“allocation”, “ci\_permuation”, “initialization”, ”initialization\_ci”,

“update\_bound”) are also available on this link. To use the “ci” function, the variable indicating the group of intervention and the cluster ID must be denoted *arm* and *cluster*, respectively, due to the way the function was coded.

We consider a dataset, denoted *dataset*, with the following variables: the observed survival time (denoted *time*), the event indicator = 1 if the observed time corresponds to the time between the origin and the occurrence of the survival event and 0 otherwise (denoted *indicator*), the trial group = 0 for the control group and 1 for the intervention group (denoted *group*) and the cluster ID (denoted *cluster*). The horizon time  $t^*$  (denoted *t\_star*) is fixed to 365.

### G.2.1. Kaplan-Meier-based method

```
t_star = 365

#function to compute the bootstrap variance
variance_bootstrap <- function(dataset, t_star) {
  cluster.id.0 <-
  unique(as.numeric(dataset[which(dataset$group == 0),
  ]$cluster))
  cluster.id.1 <-
  unique(as.numeric(dataset[which(dataset$group == 1),
  ]$cluster))

  K0 <- length(cluster.id.0)
  K1 <- length(cluster.id.1)

  #sample K0 and K1 clusters in control and intervention
  groups
  c <-c(sample(cluster.id.0, K0, replace = TRUE),
  sample(cluster.id.1, K1, replace = TRUE))

  #select the clusters in the data.frame
  data.boot <- dataset[unlist(lapply(c, function(x)
  which(dataset$cluster == x))), ]

  # Check
  while(min(max(data.boot[which(data.boot$group==1),]$time),
  max(data.boot[which(data.boot$group == 0),
  ]$time))<t_star){
```

```

cluster.id.0 <-
unique(as.numeric(dataset[which(datase$group == 0),
]$cluster))
cluster.id.1 <-
unique(as.numeric(dataset[which(dataset$group == 1),
]$cluster))

K0 <- length(cluster.id.0)
K1 <- length(cluster.id.1)

#sample K0 and K1 clusters in control and intervention
group
c <- c(sample(cluster.id.0, K0, replace=TRUE),
sample(cluster.id.1, K1, replace=TRUE))

data.boot <- dataset[unlist(lapply(c , function(x)
which(dataset$cluster == x))), ]
}

#Analyze with the bootstrap dataset
results <- rmst2(data.boot$time, data.boot$indicator,
data.boot$group, tau= t_star)

#Return the difference in RMST
return(results$unadjusted.result[1,1])
}

kaplan_meier <- rmst2(dataset$time, dataset$indicator,
dataset$group, tau = t_star)

bootstrap <- replicate(10000, variance_bootstrap(dataset,
t_star))

#Difference in RMST
kaplan_meier$unadjusted.result[1,"Est."]

#Confidence interval
quantile(bootstrap, 0.025)
quantile(bootstrap, 0.975)

G.2.2. Pseudo-values regression based methods
t_star = 365

#Computation of the pseudo-values
dataset$pv <- pseudomean(dataset$time,
dataset$indicator,

```



```

tmax=t_star)

#Exchangeable matrix
PV1 <-summary(gee(pv ~ group,
  data = dataset,
  id = cluster,
  family = gaussian,
  corstr = "exchangeable))

#Independent matrix
PV2 <- summary(gee(pv ~ group,
  data = dataset,
  id = cluster,
  family = gaussian,
  corstr = "independence"))

#Difference in RMST
PV1$coefficients["group","Estimate"]
PV2$coefficients["group","Estimate"]

#Model-based confidence interval
PV1$coefficients["group","Estimate"]-
qnorm(0.975)*sqrt(PV1$coefficients["group","Robust S.E."]^2)
PV1$coefficients["group","Estimate"]+qnorm(0.975)*sqrt(PV1$coe
fficients["group","Robust S.E."]^2)

PV2$coefficients["group","Estimate"]-
qnorm(0.975)*sqrt(PV2$coefficients["group","Robust S.E."]^2)
PV2$coefficients["group","Estimate"]+qnorm(0.975)*sqrt(PV2$coe
fficients["group","Robust S.E."]^2)

#Permutation-based confidence interval
base$arm <- base$group

ci_exc <- ci(dataset, "exchangeable")
ci_ind <- ci(dataset, "independence")

```