"A novel calibration framework for survival analysis when a binary covariate is measured at sparse time points"
Supplementary Materials

Daniel Nevo, Tsuyoshi Hamada, Shuji Ogino and Molin Wang

A Asymptotic theory and proofs

In this section, we provide proofs for the results stated in Section 5 of the main paper. We first present proofs for consistency and normality for the ordinary calibration, and then for risk-set calibration. The proofs for the RSC estimator are similar to those of the OC estimator, and are included here for completeness.

A.1 Ordinary calibration

We first show that $\hat{\boldsymbol{\theta}} \to \boldsymbol{\theta}^\star$. We repeat the notions from Section 5 of the main paper and add more notations. Denote $\nu_i^{\boldsymbol{\theta},\boldsymbol{\eta}}(t) = E_{P_{\boldsymbol{\eta}}}[\exp(\beta X_i(t))|\mathcal{G}_{it}]$ and $\nu_i^0(t) = E_{\mathcal{P}_0}[\exp(\beta^0 X_i(t))|\mathcal{G}_{it}]$, where for any $\boldsymbol{\eta}$, $E_{P_{\boldsymbol{\eta}}}$ is the expectation under the PH model for V, Equation (4.7) in the main paper, and $E_{\mathcal{P}_0}$ is the expectation under the true distribution. Denote also

$$\begin{aligned}
\boldsymbol{S}^{(m)}(\boldsymbol{\theta},\boldsymbol{\eta},t) &= \frac{1}{n}\sum_{i=1}^{n}\left[Y_i(t)\exp(\boldsymbol{\gamma}^T\boldsymbol{Z}_i)\nu_i^{\boldsymbol{\theta},\boldsymbol{\eta}}(t)\boldsymbol{a}_i(\boldsymbol{\theta},\boldsymbol{\eta},t)^{\otimes m}\right] \\
\boldsymbol{s}^{(m)}(\boldsymbol{\theta},\boldsymbol{\eta},t) &= E[\boldsymbol{S}^{(m)}(\boldsymbol{\theta},\boldsymbol{\eta},t)] \\
\overline{\boldsymbol{s}}^{(m)}(\boldsymbol{\eta},t) &= E\left[\frac{1}{n}\sum_{i=1}^{n}\left(Y_i(t)\lambda_0(t)\exp(\boldsymbol{\gamma}_0^T\boldsymbol{Z}_i)\nu_i^0(t)\boldsymbol{a}_i(\boldsymbol{\theta},\boldsymbol{\eta},t)^{\otimes m}\right)\right]
\end{aligned}$$

where $\boldsymbol{a}_i(\boldsymbol{\theta},\boldsymbol{\eta},t) = \begin{pmatrix} \frac{\exp(\beta)P_{\boldsymbol{\eta}}[X_i(t)=1|\mathcal{G}_{it}]}{1+(\exp(\beta)-1)P_{\boldsymbol{\eta}}[X_i(t)=1|\mathcal{G}_{it}]} \\ \boldsymbol{Z}_i \end{pmatrix}$ and where for any vector $\boldsymbol{x}$, $\boldsymbol{x}^{\otimes 0} = 1, \boldsymbol{x}^{\otimes 1} = \boldsymbol{x}$, and $\boldsymbol{x}^{\otimes 2} = \boldsymbol{x}\boldsymbol{x}^T$. Observe that $\lambda_0(t)\exp(\boldsymbol{\gamma}_0^T\boldsymbol{Z}_i)\nu_i^0(t)$ is the true hazard function (conditionally on $\mathcal{G}_{it}$). Define also $\boldsymbol{u}_{\boldsymbol{\theta}}^{\mathcal{G}}(\boldsymbol{\theta};\boldsymbol{\eta}) = \int_0^\tau \overline{\boldsymbol{s}}^{(1)}(\boldsymbol{\eta},t)dt - \int_0^\tau \frac{\boldsymbol{s}^{(1)}(\boldsymbol{\theta},\boldsymbol{\eta},t)}{s^{(0)}(\boldsymbol{\theta},\boldsymbol{\eta},t)}\overline{s}^{(0)}(\boldsymbol{\eta},t)dt$,

and let $\boldsymbol{I}^{\mathcal{G}}_{\boldsymbol{\theta}}(\boldsymbol{\theta},\boldsymbol{\eta}) = \nabla_{\boldsymbol{\theta}}\boldsymbol{u}^{\mathcal{G}}_{\boldsymbol{\theta}}(\boldsymbol{\theta};\boldsymbol{\eta})$ and $\boldsymbol{I}^{\mathcal{G}}_{\boldsymbol{\eta}}(\boldsymbol{\theta},\boldsymbol{\eta}) = \nabla_{\boldsymbol{\eta}}\boldsymbol{u}^{\mathcal{G}}_{\boldsymbol{\theta}}(\boldsymbol{\theta};\boldsymbol{\eta})$. We impose the following standard regularity assumptions:

(A1) The number of knots K does not grow with the sample size n.

(A2) $\hat{\boldsymbol{\eta}} \xrightarrow{p} \boldsymbol{\eta}^{\star}$, for some $\boldsymbol{\eta}^{\star}$.

(A3) $\boldsymbol{s}^{(m)}(\boldsymbol{\theta},\boldsymbol{\eta},t), m = 0,1,2$ are continuous and bounded functions of $\boldsymbol{\theta}$, for any $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$ in the neighborhoods Θ and $\mathcal{H}$ of $\boldsymbol{\theta}^{\star}$ and $\boldsymbol{\eta}^{\star}$, respectively, and for all $t \in [0,\tau]$. Furthermore, $s^{(0)}(\boldsymbol{\theta}^{\star},\boldsymbol{\eta}^{\star},t)$ is bounded away from zero.

(A4) The components of $\boldsymbol{Z}_i$ are bounded for all i.

(A5) The matrix $\boldsymbol{I}^{\mathcal{G}}_{\boldsymbol{\theta}}(\boldsymbol{\theta},\boldsymbol{\eta}^{\star})$ is continuous in $\boldsymbol{\theta}$ and positive definite at $\boldsymbol{\theta}^{\star}$.

Assumption (A1) could be relaxed, see Wang et al. (2016). But for simplicity, we consider (A1) as given above. By the Weak Law of Large Numbers,

$\sup_{t\in[0,\tau],\boldsymbol{\theta}\in\Theta,\boldsymbol{\eta}\in\mathcal{H}} |\boldsymbol{S}^{(m)}(\boldsymbol{\theta},\boldsymbol{\eta},t) - \boldsymbol{s}^{(m)}(\boldsymbol{\theta},\boldsymbol{\eta},t)| \xrightarrow{p} 0.$

By the assumptions above and using arguments similar to those of Andersen & Gill (1982) as implemented by Lin & Wei (1989), it follows that for any $\boldsymbol{\theta} \in \Theta$ and $\boldsymbol{\eta} \in \mathcal{H}$, $\boldsymbol{U}^{\mathcal{G}}_{\boldsymbol{\theta}}(\boldsymbol{\theta};\boldsymbol{\eta}) \xrightarrow{p} \boldsymbol{u}^{\mathcal{G}}_{\boldsymbol{\theta}}(\boldsymbol{\theta};\boldsymbol{\eta})$. Recall that $\hat{\boldsymbol{\theta}}$ is the solution of $\boldsymbol{U}^{\mathcal{G}}(\boldsymbol{\theta};\hat{\boldsymbol{\eta}}) = 0$ and observe that

$$\boldsymbol{U}^{\mathcal{G}}_{\boldsymbol{\theta}}(\boldsymbol{\theta};\hat{\boldsymbol{\eta}}) = \boldsymbol{U}^{\mathcal{G}}_{\boldsymbol{\theta}}(\boldsymbol{\theta};\boldsymbol{\eta}^{\star}) + (\boldsymbol{U}^{\mathcal{G}}_{\boldsymbol{\theta}}(\boldsymbol{\theta};\hat{\boldsymbol{\eta}}) - \boldsymbol{U}^{\mathcal{G}}_{\boldsymbol{\theta}}(\boldsymbol{\theta};\boldsymbol{\eta}^{\star})) = \boldsymbol{U}^{\mathcal{G}}_{\boldsymbol{\theta}}(\boldsymbol{\theta};\boldsymbol{\eta}^{\star}) + o_p(1)$$

where the last equality holds by Assumption (A2) and since $\boldsymbol{a}_i$ is bounded for finite values of β. Let $\boldsymbol{\theta}^{\star}$ be the solution of $\boldsymbol{u}^{\mathcal{G}}_{\boldsymbol{\theta}}(\boldsymbol{\theta};\boldsymbol{\eta}^{\star}) = 0$. By the assumptions above, and specifically Assumption (A5), $\hat{\boldsymbol{\theta}} \xrightarrow{p} \boldsymbol{\theta}^{\star}$. In particular, $\hat{\beta} \xrightarrow{p} \beta^{\star}$.

Regrading asymptotic normally, by a Taylor expansion

$$0 = \boldsymbol{U}^{\mathcal{G}}_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}};\hat{\boldsymbol{\eta}}) = \boldsymbol{U}^{\mathcal{G}}_{\boldsymbol{\theta}}(\boldsymbol{\theta}^{\star};\boldsymbol{\eta}^{\star}) + [\nabla_{\boldsymbol{\theta}}\boldsymbol{U}^{\mathcal{G}}_{\boldsymbol{\theta}}(\boldsymbol{\theta}^{\star};\boldsymbol{\eta}^{\star})](\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^{\star}) + [\nabla_{\boldsymbol{\eta}}\boldsymbol{U}^{\mathcal{G}}_{\boldsymbol{\theta}}(\boldsymbol{\theta}^{\star};\boldsymbol{\eta}^{\star})](\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}^{\star}) + o_p(n^{-1/2})$$

which can be rearranged as

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^{\star}) = [-\nabla_{\boldsymbol{\theta}}\boldsymbol{U}^{\mathcal{G}}_{\boldsymbol{\theta}}(\boldsymbol{\theta}^{\star};\boldsymbol{\eta}^{\star})]^{-1}\sqrt{n}\{\boldsymbol{U}^{\mathcal{G}}_{\boldsymbol{\theta}}(\boldsymbol{\theta}^{\star};\boldsymbol{\eta}^{\star}) + [\nabla_{\boldsymbol{\eta}}\boldsymbol{U}^{\mathcal{G}}_{\boldsymbol{\theta}}(\boldsymbol{\theta}^{\star};\boldsymbol{\eta}^{\star})(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}^{\star})]\} + o_p(1). \tag{1}$$

Similarly to Andersen & Gill (1982) and Lin & Wei (1989), by the assumptions given above, $\nabla_{\boldsymbol{\theta}}\boldsymbol{U}^{\mathcal{G}}_{\boldsymbol{\theta}}(\boldsymbol{\theta}^{\star};\boldsymbol{\eta}^{\star}) \xrightarrow{p} \boldsymbol{I}^{\mathcal{G}}_{\boldsymbol{\theta}}(\boldsymbol{\theta}^{\star},\boldsymbol{\eta}^{\star})$, and by invoking similar arguments $\nabla_{\boldsymbol{\eta}}\boldsymbol{U}^{\mathcal{G}}_{\boldsymbol{\theta}}(\boldsymbol{\theta}^{\star};\boldsymbol{\eta}^{\star}) \xrightarrow{p} \boldsymbol{I}^{\mathcal{G}}_{\boldsymbol{\eta}}(\boldsymbol{\theta}^{\star},\boldsymbol{\eta}^{\star})$. Let $\mathcal{N}(t) = E[n^{-1}\sum_{i=1}^{n} N_i(t)]$. By arguments similar to those of Lin & Wei (1989), it can be shown that $n^{1/2}\boldsymbol{U}^{\mathcal{G}}_{\boldsymbol{\theta}}(\boldsymbol{\theta}^{\star};\boldsymbol{\eta}^{\star}) = n^{-1/2}\sum_{i=1}^{n}\boldsymbol{b}_i(\boldsymbol{\theta}^{\star};\boldsymbol{\eta}^{\star}) + o_p(1)$, where

$$\begin{aligned}\boldsymbol{b}_i(\boldsymbol{\theta};\boldsymbol{\eta}) = &\int_0^{\tau}\left[\boldsymbol{a}_i(\boldsymbol{\theta},\boldsymbol{\eta},t) - \frac{\boldsymbol{s}^{(1)}(\boldsymbol{\theta},\boldsymbol{\eta},t)}{s^{(0)}(\boldsymbol{\theta},\boldsymbol{\eta},t)}\right] dN_i(t) \\ &- \int_0^{\infty}\frac{Y_i(t)\exp(\boldsymbol{\gamma}^T\boldsymbol{Z}_i)\nu_i^{\boldsymbol{\theta},\boldsymbol{\eta}}(t)}{s^{(0)}(\boldsymbol{\theta},\boldsymbol{\eta},t)}\left[\boldsymbol{a}_i(\boldsymbol{\theta},\boldsymbol{\eta},t) - \frac{\boldsymbol{s}^{(1)}(\boldsymbol{\theta},\boldsymbol{\eta},t)}{s^{(0)}(\boldsymbol{\theta},\boldsymbol{\eta},t)}\right] d\mathcal{N}(t) + o_p(1).\end{aligned}$$

Regarding $(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}^\star)$, by Assumption (A1), it is a finite-size vector, and as explained in Wang et al. (2016), it can be treated with the standard tools for parametric models. Denote $\ell^V = \log L^V$. By further incorporating standard theory of misspecified likelihood-based models and Assumption (A2), we may write

$$\sqrt{n}(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}^\star) = -\left(\frac{1}{n}\sum_{i=1}^{n} \nabla_{\boldsymbol{\eta}\boldsymbol{\eta}} \ell_i^V(\boldsymbol{\eta}^\star)\right)^{-1} \left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} \nabla_{\boldsymbol{\eta}} \ell_i^V(\boldsymbol{\eta}^\star)\right) + o_p(1).$$

Substituting this in (1), we get

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^\star) = [-\nabla_{\boldsymbol{\theta}} \boldsymbol{U}_{\boldsymbol{\theta}}^{\mathcal{G}}(\boldsymbol{\theta}^\star; \boldsymbol{\eta}^\star)]^{-1} \left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} \boldsymbol{r}_i(\boldsymbol{\theta}^\star, \boldsymbol{\eta}^\star)\right) + o_p(1)$$

where

$$\boldsymbol{r}_i(\boldsymbol{\theta}, \boldsymbol{\eta}) = \boldsymbol{b}_i(\boldsymbol{\theta}, \boldsymbol{\eta}) - [\nabla_{\boldsymbol{\eta}} \boldsymbol{U}_{\boldsymbol{\theta}}^{\mathcal{G}}(\boldsymbol{\theta}; \boldsymbol{\eta})] \left(\frac{1}{n}\sum_{j=1}^{n} \nabla_{\boldsymbol{\eta}\boldsymbol{\eta}} \ell_j^V(\boldsymbol{\eta})\right)^{-1} \nabla_{\boldsymbol{\eta}} \ell_i^V(\boldsymbol{\eta}).$$

Therefore, by the multivariate central limit theorem and Slutsky's theorem, we conclude that $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^\star)$ is asymptotically normally distributed with covariance matrix

$$[\boldsymbol{I}_{\boldsymbol{\theta}}^{\mathcal{G}}(\boldsymbol{\theta}^\star, \boldsymbol{\eta}^\star)]^{-1} E(\boldsymbol{r}_i(\boldsymbol{\theta}^\star, \boldsymbol{\eta}^\star)^{\otimes 2}) [\boldsymbol{I}_{\boldsymbol{\theta}}^{\mathcal{G}}(\boldsymbol{\theta}^\star, \boldsymbol{\eta}^\star)]^{-1}$$

which can be consistently estimated by $\hat{\mathcal{V}}$ in Equation (5.8) in the main text, by replacing parameters with their estimates and expectations with their corresponding sample version.

A.2 Risk-set calibration

We start by introducing additional relevant notations. Let $t_1^\star, ..., t_H^\star$ be the ordered H event times observed in the data and let $\tilde{\boldsymbol{\eta}} = (\tilde{\boldsymbol{\eta}}(t_1), \tilde{\boldsymbol{\eta}}(t_2), ...\tilde{\boldsymbol{\eta}}(t_H))$ be the vector of the time-dependent calibration parameters. At each observed event time $t^\star$, $\tilde{\boldsymbol{\eta}}(t^\star)$ is estimated by maximizing the log-likelihood $\sum_{i=1}^{n} Y_i(t^\star) \ell_i^V(\tilde{\boldsymbol{\eta}}(t^\star))$. Denote also $\nu_i^{\mathcal{F},\boldsymbol{\theta},\tilde{\boldsymbol{\eta}}}(t) = E_{P_{\tilde{\boldsymbol{\eta}}}}[\exp(\beta X_i(t))|\mathcal{F}_{it}]$ and $\nu_i^{\mathcal{F},0}(t) = E_{\mathcal{P}_0}[\exp(\beta^0 X_i(t))|\mathcal{F}_{it}]$ which are analogous to $\nu_i^{\boldsymbol{\theta},\boldsymbol{\eta}}(t)$ and $\nu_i^0(t)$ in the theory for the OC estimator.

The RSC estimator maximizes $L^{\mathcal{F}}(\boldsymbol{\theta}, \hat{\tilde{\boldsymbol{\eta}}})$, or alternatively, solves the estimating equation $U^{\mathcal{F}}(\boldsymbol{\theta}, \hat{\tilde{\boldsymbol{\eta}}}) = 0$ where

$$\boldsymbol{U}^{\mathcal{F}}(\boldsymbol{\theta}; \tilde{\boldsymbol{\eta}}) = \frac{1}{n}\sum_{i=1}^{n} \int_0^\tau \left[\tilde{\boldsymbol{a}}_i(\boldsymbol{\theta}, \tilde{\boldsymbol{\eta}}(t), t) - \frac{\tilde{\boldsymbol{S}}^{(1)}(\boldsymbol{\theta}, \tilde{\boldsymbol{\eta}}(t), t)}{\tilde{S}^{(0)}(\boldsymbol{\theta}, \tilde{\boldsymbol{\eta}}(t), t)}\right] dN_i(t)$$

with

$$\tilde{\boldsymbol{S}}^{(m)}(\boldsymbol{\theta},\tilde{\boldsymbol{\eta}}(t),t)=\frac{1}{n}\sum_{i=1}^{n}\left[Y_i(t)\exp(\boldsymbol{\gamma}^T\boldsymbol{Z}_i)\nu_i^{\mathcal{F},\boldsymbol{\theta},\tilde{\boldsymbol{\eta}}(t)}(t)\tilde{\boldsymbol{a}}_i(\boldsymbol{\theta},\tilde{\boldsymbol{\eta}}(t),t)^{\otimes m}\right]$$

$$\tilde{\boldsymbol{s}}^{(m)}(\boldsymbol{\theta},\tilde{\boldsymbol{\eta}}(t),t)=E[\tilde{\boldsymbol{S}}^{(m)}(\boldsymbol{\theta},\tilde{\boldsymbol{\eta}}(t),t)]$$

$$\tilde{\bar{\boldsymbol{s}}}^{(m)}(\tilde{\boldsymbol{\eta}}(t),t)=E\left[\frac{1}{n}\sum_{i=1}^{n}\left(Y_i(t)\lambda_0(t)\exp(\boldsymbol{\gamma}_0^T\boldsymbol{Z}_i)\nu_i^{\mathcal{F},0}(t)\tilde{\boldsymbol{a}}_i(\boldsymbol{\theta},\tilde{\boldsymbol{\eta}}(t),t)^{\otimes m}\right)\right]$$

and $\tilde{\boldsymbol{a}}_i(\boldsymbol{\theta},\tilde{\boldsymbol{\eta}}(t),t)=\begin{pmatrix}\frac{\exp(\beta)P_{\tilde{\boldsymbol{\eta}}(t)}[X_i(t)=1|\mathcal{F}_{it}]}{1+(\exp(\beta)-1)P_{\tilde{\boldsymbol{\eta}}(t)}[X_i(t)=1|\mathcal{F}_{it}]}\\ \boldsymbol{Z}_i\end{pmatrix}$. Similarly to the OC case, $\lambda_0(t)\exp(\boldsymbol{\gamma}_0^T\boldsymbol{Z}_i)\nu_i^{\mathcal{F},0}(t)$ is the true hazard function (conditionally on $\mathcal{F}_{it}$). Define also $\boldsymbol{u}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta};\tilde{\boldsymbol{\eta}})=\int_0^\tau\tilde{\boldsymbol{s}}^{(1)}(\tilde{\boldsymbol{\eta}}(t),t)dt-\int_0^\tau\frac{\tilde{\boldsymbol{s}}^{(1)}(\boldsymbol{\theta},\tilde{\boldsymbol{\eta}}(t),t)}{\tilde{s}^{(0)}(\boldsymbol{\theta},\tilde{\boldsymbol{\eta}}(t),t)}\tilde{s}^{(0)}(\tilde{\boldsymbol{\eta}}(t),t)dt$, and let $\boldsymbol{I}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta},\tilde{\boldsymbol{\eta}})=\nabla_{\boldsymbol{\theta}}\boldsymbol{u}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta};\tilde{\boldsymbol{\eta}})$ and $\boldsymbol{I}_{\tilde{\boldsymbol{\eta}}}^{\mathcal{F}}(\boldsymbol{\theta},\tilde{\boldsymbol{\eta}})=\nabla_{\tilde{\boldsymbol{\eta}}}\boldsymbol{u}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta};\tilde{\boldsymbol{\eta}})$.
The regularity assumptions are slightly adjusted:

($\widetilde{\text{A}}1$) The number of knots K does not grow with the sample size n.

($\widetilde{\text{A}}2$) $\hat{\tilde{\boldsymbol{\eta}}}\xrightarrow{p}\tilde{\boldsymbol{\eta}}^\star$, for some $\tilde{\boldsymbol{\eta}}^\star$.

($\widetilde{\text{A}}3$) $\tilde{\boldsymbol{s}}^{(m)}(\boldsymbol{\theta},\tilde{\boldsymbol{\eta}}(t),t),m=0,1,2$ are continuous and bounded functions of $\boldsymbol{\theta}$, for any $\boldsymbol{\theta}$ and $\tilde{\boldsymbol{\eta}}$ in the neighborhoods Θ and $\tilde{\mathcal{H}}$ of $\boldsymbol{\theta}^\star$ and $\tilde{\boldsymbol{\eta}}^\star$, respectively, and for all $t\in[0,\tau]$. Furthermore, $\tilde{s}^{(0)}(\boldsymbol{\theta}^\star,\boldsymbol{\eta}^\star,t)$ is bounded away from zero.

($\widetilde{\text{A}}4$) The components of $\boldsymbol{Z}_i$ are bounded for all i.

($\widetilde{\text{A}}5$) The matrix $\boldsymbol{I}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta},\tilde{\boldsymbol{\eta}}^\star)$ is continuous in $\boldsymbol{\theta}$ and positive definite at $\boldsymbol{\theta}^\star$.

As before, Assumption ($\tilde{\text{A}}1$) could be relaxed, see Wang et al. (2016). By the Weak Law of Large Numbers,

$$\sup_{t\in[0,\tau],\boldsymbol{\theta}\in\Theta,\boldsymbol{\eta}\in\mathcal{H}}|\tilde{\boldsymbol{S}}^{(m)}(\boldsymbol{\theta},\tilde{\boldsymbol{\eta}}(t),t)-\tilde{\boldsymbol{s}}^{(m)}(\boldsymbol{\theta},\tilde{\boldsymbol{\eta}}(t),t)|\xrightarrow{p}0.$$

By the assumptions above and using arguments similar to those of Andersen & Gill (1982) as implemented by Lin & Wei (1989), it follows that for any $\boldsymbol{\theta}\in\Theta$ and $\tilde{\boldsymbol{\eta}}\in\tilde{\mathcal{H}}$, $\boldsymbol{U}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta};\tilde{\boldsymbol{\eta}})\xrightarrow{p}\boldsymbol{u}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta};\tilde{\boldsymbol{\eta}})$. Recall that $\hat{\boldsymbol{\theta}}$ is the solution of $\boldsymbol{U}^{\mathcal{F}}(\boldsymbol{\theta};\hat{\tilde{\boldsymbol{\eta}}})=0$ and observe that

$$\boldsymbol{U}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta};\hat{\tilde{\boldsymbol{\eta}}})=\boldsymbol{U}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta};\tilde{\boldsymbol{\eta}}^\star)+(\boldsymbol{U}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta};\hat{\tilde{\boldsymbol{\eta}}})-\boldsymbol{U}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta};\tilde{\boldsymbol{\eta}}^\star))=\boldsymbol{U}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta};\tilde{\boldsymbol{\eta}}^\star)+o_p(1)$$

where the last equality holds by Assumption ($\widetilde{\text{A}}2$) and since $\tilde{\boldsymbol{a}}_i$ is bounded for finite values of β. Let $\boldsymbol{\theta}^\star$ be the solution of $\boldsymbol{u}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta};\tilde{\boldsymbol{\eta}}^\star)=0$. By the assumptions above, and specifically Assumption ($\widetilde{\text{A}}5$), $\hat{\boldsymbol{\theta}}\xrightarrow{p}\boldsymbol{\theta}^\star$. In particular, $\hat{\beta}\xrightarrow{p}\beta^\star$.

Regrading asymptotic normally, by a Taylor expansion

$$0 = \boldsymbol{U}_{\boldsymbol{\theta}}^{\mathcal{F}}(\hat{\boldsymbol{\theta}}; \hat{\tilde{\boldsymbol{\eta}}}) = \boldsymbol{U}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta}^\star; \tilde{\boldsymbol{\eta}}^\star) + [\nabla_{\boldsymbol{\theta}} \boldsymbol{U}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta}^\star; \tilde{\boldsymbol{\eta}}^\star)](\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^\star) + [\nabla_{\tilde{\boldsymbol{\eta}}} \boldsymbol{U}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta}^\star; \tilde{\boldsymbol{\eta}}^\star)](\hat{\tilde{\boldsymbol{\eta}}} - \tilde{\boldsymbol{\eta}}^\star) + o_p(n^{-1/2})$$

which can be rearranged as

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^\star) = [-\nabla_{\boldsymbol{\theta}} \boldsymbol{U}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta}^\star; \tilde{\boldsymbol{\eta}}^\star)]^{-1} \sqrt{n} \{ \boldsymbol{U}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta}^\star; \tilde{\boldsymbol{\eta}}^\star) + [\nabla_{\tilde{\boldsymbol{\eta}}} \boldsymbol{U}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta}^\star; \tilde{\boldsymbol{\eta}}^\star)(\hat{\tilde{\boldsymbol{\eta}}} - \tilde{\boldsymbol{\eta}}^\star)] \} + o_p(1). \tag{2}$$

Similarly to Andersen & Gill (1982) and Lin & Wei (1989), by the assumptions given above, $\nabla_{\boldsymbol{\theta}} \boldsymbol{U}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta}^\star; \tilde{\boldsymbol{\eta}}^\star) \xrightarrow{p} \boldsymbol{I}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta}^\star, \tilde{\boldsymbol{\eta}}^\star)$, and by invoking similar arguments $\nabla_{\tilde{\boldsymbol{\eta}}} \boldsymbol{U}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta}^\star; \tilde{\boldsymbol{\eta}}^\star) \xrightarrow{p} \boldsymbol{I}_{\tilde{\boldsymbol{\eta}}}^{\mathcal{F}}(\boldsymbol{\theta}^\star, \tilde{\boldsymbol{\eta}}^\star)$. Let $\mathcal{N}(t) = E[n^{-1} \sum_{i=1}^n N_i(t)]$. By arguments similar to those of Lin & Wei (1989), it can be shown that $n^{1/2} \boldsymbol{U}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta}^\star; \tilde{\boldsymbol{\eta}}^\star) = n^{-1/2} \sum_{i=1}^n \boldsymbol{b}_i(\boldsymbol{\theta}^\star; \tilde{\boldsymbol{\eta}}^\star) + o_p(1)$, where

$$\begin{aligned}
\boldsymbol{b}_i(\boldsymbol{\theta}; \tilde{\boldsymbol{\eta}}) = & \int_0^\tau \left[\tilde{\boldsymbol{a}}_i(\boldsymbol{\theta}, \tilde{\boldsymbol{\eta}}(t), t) - \frac{\tilde{\boldsymbol{s}}^{(1)}(\boldsymbol{\theta}, \tilde{\boldsymbol{\eta}}(t), t)}{\tilde{s}^{(0)}(\boldsymbol{\theta}, \tilde{\boldsymbol{\eta}}(t), t)} \right] dN_i(t) \\
& - \int_0^\infty \frac{Y_i(t) \exp(\boldsymbol{\gamma}^T \boldsymbol{Z}_i) \nu_i^{\mathcal{F}, \boldsymbol{\theta}, \tilde{\boldsymbol{\eta}}}(t)}{\tilde{s}^{(0)}(\boldsymbol{\theta}, \tilde{\boldsymbol{\eta}}(t), t)} \left[\tilde{\boldsymbol{a}}_i(\boldsymbol{\theta}, \tilde{\boldsymbol{\eta}}(t), t) - \frac{\tilde{\boldsymbol{s}}^{(1)}(\boldsymbol{\theta}, \tilde{\boldsymbol{\eta}}(t), t)}{\tilde{s}^{(0)}(\boldsymbol{\theta}, \tilde{\boldsymbol{\eta}}(t), t)} \right] d\mathcal{N}(t) + o_p(1).
\end{aligned}$$

Regarding $(\hat{\tilde{\boldsymbol{\eta}}} - \tilde{\boldsymbol{\eta}}^\star)$, by Assumption ($\widetilde{\text{A}}1$), it is a finite-size vector, and as explained in Wang et al. (2016), it can be treated with the standard tools for parametric models. Denote $\ell^V = \log L^V$. By further incorporating standard theory of misspecified likelihood-based models and Assumption ($\widetilde{\text{A}}2$), we may write

$$\sqrt{n}(\hat{\tilde{\boldsymbol{\eta}}} - \tilde{\boldsymbol{\eta}}^\star) = - \left(\frac{1}{n} \sum_{i=1}^n \nabla_{\tilde{\boldsymbol{\eta}}\tilde{\boldsymbol{\eta}}} \ell_i^V(\tilde{\boldsymbol{\eta}}^\star) \right)^{-1} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n \nabla_{\tilde{\boldsymbol{\eta}}} \ell_i^V(\tilde{\boldsymbol{\eta}}^\star) \right) + o_p(1).$$

Substituting this into (2), we get

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^\star) = [-\nabla_{\boldsymbol{\theta}} \boldsymbol{U}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta}^\star; \tilde{\boldsymbol{\eta}}^\star)]^{-1} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n \boldsymbol{r}_i(\boldsymbol{\theta}^\star, \tilde{\boldsymbol{\eta}}^\star) \right) + o_p(1)$$

where

$$\boldsymbol{r}_i(\boldsymbol{\theta}, \tilde{\boldsymbol{\eta}}) = \boldsymbol{b}_i(\boldsymbol{\theta}, \tilde{\boldsymbol{\eta}}) - [\nabla_{\tilde{\boldsymbol{\eta}}} \boldsymbol{U}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta}; \tilde{\boldsymbol{\eta}})] \left(\frac{1}{n} \sum_{j=1}^n \nabla_{\tilde{\boldsymbol{\eta}}\tilde{\boldsymbol{\eta}}} \ell_j^V(\tilde{\boldsymbol{\eta}}) \right)^{-1} \nabla_{\tilde{\boldsymbol{\eta}}} \ell_i^V(\tilde{\boldsymbol{\eta}}).$$

Therefore, by the multivariate central limit theorem and Slutsky's theorem, we conclude that $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^\star)$ is asymptotically normally distributed with covariance matrix

$$[\boldsymbol{I}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta}^\star, \tilde{\boldsymbol{\eta}}^\star)]^{-1} E(\boldsymbol{r}_i(\boldsymbol{\theta}^\star, \tilde{\boldsymbol{\eta}}^\star)^{\otimes 2}) [\boldsymbol{I}_{\boldsymbol{\theta}}^{\mathcal{F}}(\boldsymbol{\theta}^\star, \tilde{\boldsymbol{\eta}}^\star)]^{-1}$$

which can be consistently estimated by

$$\widehat{\widetilde{\mathcal{V}}} = [-\nabla_{\boldsymbol{\theta}} \boldsymbol{U}^{\mathcal{F}}_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}_{RSC}, \hat{\tilde{\boldsymbol{\eta}}})]^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} \hat{\boldsymbol{r}}_i(\hat{\boldsymbol{\theta}}_{RSC}, \hat{\tilde{\boldsymbol{\eta}}}) \hat{\boldsymbol{r}}_i^T(\hat{\boldsymbol{\theta}}_{RSC}, \hat{\tilde{\boldsymbol{\eta}}}) \right) [-\nabla_{\boldsymbol{\theta}} \boldsymbol{U}^{\mathcal{G}}_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}_{RSC}, \hat{\tilde{\boldsymbol{\eta}}})]^{-1}$$

and replacing parameters with their estimates and expectations with their corresponding sample version.

B Description of the R package `ICcalib`

In this section, we describe the `ICcalib` **R** package and how to use it. Details on function arguments, returns and syntax are given in the package manual (CRAN or Github). The functions and scripts that were used in the simulation studies and data analysis demonstrate how to use the package. In its current version (version `1.0.5`) the `ICcalib` carries out the analysis by the following steps:

1. Fit a calibration model from the interval-censored data about the binary exposure. The names of these functions start with `FitCalib` and continues with the name of the model.

   - `FitCalibCox` fits a PH calibration model (with covariates).
   - `FitCalibWeibull` fits a Weibull calibration model.
   - `FitCalibNPMLE` fits a non-parametric calibration model.

2. Calculate $\widehat{P}(X_i(t) = 1|\mathcal{G}_{it})$ for all $i = 1, ..., n$ and all t for which a main event was observed. This step is carried out by looping over the event times, and use one of the following functions to calculate $\widehat{P}(X_i(t) = 1|\mathcal{G}_{it})$.

   - `CalcCoxCalibP` calculates the probability using the results of a PH calibration model fitting (with covariates).
   - `CalcWeibullCalibP` calculates the probability using the results of a Weibull calibration model fitting.
   - `CalcNpmleCalibP` calculates the probability using the results of a non-parametric calibration model fitting.

3. Obtain estimates by maximizing the partial likelihood given in Equation (4.3) of the main paper, using the functions `CoxLogLikX` or `CoxLogLik`, the former when there are no additional covariates in the main model (i.e., no $\boldsymbol{Z}$) and the latter in the presence of such covariates.

4. For variance calculation and inference:

- For PH calibration models, first the function `CalcCoxCalibPderiv` is called in a similar manner to `CalcCoxCalibP`. Then, the function `CalcVarParam` calculates the covariance matrix.
- For Weibull calibration models, first the functions `CalcWeibullCalibPderivShape` and `CalcWeibullCalibPderivScale` are called. Then, the function `CalcVarThetaWeib` returns the covariance matrix.
- For non-parametric calibration models the bootstrap is used for inference.

Similar procedures can be implemented with modified functions for risk-set calibrations. Please see the package manual and the reproducibility repository ICcalibReproduce for details and examples.

C Additional simulation studies

In this section, we present additional results of the simulation study discussed in the main paper, and description and results of further simulation studies, for when there are no baseline covariates, and the calibration model is fitted nonparametrically or parametrically. Table A.2 presents results of the study described in the paper, under additional scenarios. It extends Table 1 in the main paper by (a) including the MidI method, (b) considering additional values for β ($\log(1/7)$, $\log(1/5)$ and $\log(1/2)$ and (c) considering a larger number of potential questionnaire time points ($M^\star = 10$).

We also considered a simulation study when the true distribution of V does not depend on any covariates. Two scenarios were examined: Weibull distribution, and piecewise exponential distribution. The exact parameter values that were used can be found in the accompanied **R** code. We compared between the following methods: LVCF, MidI, parametric Weibull calibration (WB-OC) and risk-set calibration models (WB-RSC), and nonparametric calibration (NP-OC) and risk-set calibration models (NP-RSC). Under the piecewise exponential scenario, using Weibull calibration (or risk-set calibration) model results in misspecification of the calibration model. We considered the following values for $\exp(\beta^0)$: 0.2, 0.5, 1, 2, 5 and $M^\star = 2, 5, 10$ similar to the main simulation study described in the paper. For the Weibull-based methods, variance was estimated by the asymptotic variance formula with plugged-in estimates and confidence intervals were calculated according to the asymptotic normal distribution. For the nonparametric calibration and risk-calibration models, the bootstrap (with 200 iterations) was used for variance estimation and construction of confidence intervals. Values of $\exp(\hat{\beta})$ larger than 150 or smaller than 150^{-1} were excluded from the summary, details are available in the **R** code.

The results, presented in Table A.3, generally agreed with the main simulation study, which had a PH model for V. In all methods, the bias increased as β^0 was getting away from zero. The MidI method was the only method to produce biased estimates when $\beta^0 = 0$. For all β^0 values, The bias was attenuated for large $M^\star$

values. The bias of the naive methods was more substantial than the bias of the calibration methods. For large values of $|\beta^0|$, the risk-set calibration methods were less biased than the ordinary calibration methods. The nonparametric calibration methods performed close to the Weibull-based methods under Weibull distribution for V, and showed some improvement of the Weibull calibration methods under piecewise exponential distribution. The difference between the results presented in Tables 1 and A.2 compared to the results in Table A.3 is that the former present the results when the data were simulated with the distribution of V depending on covariates and the latter presents the results when data were simulated without any covariates affecting the distribution of V (i.e., $\boldsymbol{\psi} = 0$).

D Analysis of CMV data

We thank Dr. Dianne Finkelstein for making the data available and direct us where it could be found. The data were made available by Finkelstein et al. (2002), and downloaded as a zip file named `interval_censr_data.zip` from `http://hedwig.mgh.harvard.edu/biostatistics/node/32`. The `cmvshedDN.sas` read the data and export it into a `cmvshed.csv` file. The included SAS file `cmvshedDN.sas` also contains the description of the variables in the dataset.

The R file `cmvshed.R` contains our data analysis. Goggins et al. (1999) reported results based on $N = 212$ and 38 events, i.e., 38 diagnosed CMV cases. We identified $n = 221$ observations and 37 events in the data, after following exclusion criteria described in `cmvshed.R`. Goggins et al. (1999) excluded participants without at least one urine or blood samples. Our method allows to include these trial participant in the main model. They do not contribute for the calibration model fitting (they all have V censored by $[0, \infty)$).

CMV shedding can occur in either blood or urine. Therefore, $X(t)$ could be defined either using the blood shedding or the urine shedding, which do not occur at the same time. Following Goggins et al. (1999), we present two separate analyses, namely for $X(t)$ defined as CMV urine shedding and CMV blood shedding.

The available data contain only $[w_{iL}, w_{iR})$ for each study participant and unlike the aspirin and CRC data, it does not include the time points in which measurements were taken before W_{iL}. Goggins et al. (1999) reported that that urine samples were taken every 4 weeks and blood samples were taken every 12 weeks. However, we have decided not to extrapolate beyond the available data, and to use only the data in the dataset, namely $[w_{iL}, w_{iR})$.

Table A.4 presents the results of LVCF, MidI, non-parametric calibration (NP-OC), non-parametric risk-set calibration (NP-RSC) and the results reported by Goggins et al. (1999). The estimated effects are strong, even though the number of events is low to moderate. We observed a divergence between the OC and RSC estimates, especially for blood shedding exposure. The confidence intervals (CI) and standard errors (SE) were calculated by the bootstrap for the NP-OC and NP-RSC, and by standard software, ignoring the imputation, for LVCF and MidI. The confidence intervals are quite wide, as one may obtain for HR of strong effects dataset with moderate sample size. The standard errors reported by Goggins et al. (1999) are lower than the estimated standard errors in the data we used, even for the LVCF and MidI methods.

References

Andersen, P. K. & Gill, R. D. (1982), 'Cox's regression model for counting processes: a large sample study', *The annals of statistics* pp. 1100–1120.

Finkelstein, D. M., Goggins, W. B. & Schoenfeld, D. A. (2002), 'Analysis of failure time data with dependent interval censoring', *Biometrics* **58**(2), 298–304.

Goggins, W. B., Finkelstein, D. M. & Zaslavsky, A. M. (1999), 'Applying the cox proportional hazards model when the change time of a binary time-varying covariate is interval censored', *Biometrics* **55**(2), 445–451.

Lin, D. Y. & Wei, L.-J. (1989), 'The robust inference for the cox proportional hazards model', *Journal of the American statistical Association* **84**(408), 1074–1078.

Wang, L., McMahan, C. S., Hudgens, M. G. & Qureshi, Z. P. (2016), 'A flexible, computationally efficient method for fitting the proportional hazards model to interval-censored data', *Biometrics* **72**(1), 222–231.

Table A.1: Summary of the main variables in the dataset

	All Data	low CD274	PIK3CA	PTGS2
n (No. Events)	1371 (249)	278 (50)	171 (28)	672 (125)
Age at diagnosis: Mean (SD)	69.4 (9.0)	69.4 (9.0)	69.9 (9.0)	68 (8.7)
CRC Stage				
I	375 (27%)	74 (27%)	55 (32%)	177 (26%)
II	453 (33%)	99 (36%)	62 (36%)	218 (32%)
III	387 (28%)	74 (27%)	44 (26%)	199 (30%)
Missing	156 (11%)	31 (11%)	10 (6%)	78 (12%)
Pre-diagnosis Aspirin Status				
Taking	587 (43%)	118 (42%)	74 (43%)	272 (40%)
Non-taking	784 (57%)	160 (58%)	97 (57%)	400 (60%)
No. Available Questionnaires: Mean (SD)	3 (1.7)	3.1 (1.7)	3.2 (1.7)	3.3 (1.7)
No. Participants with no Questionnaires	113	21	11	47

Table A.2: . Methods compared are LVCF, MidI, PH calibration model (PH-OC) and PH risk-set calibration models (PH-RSC). The table presents mean estimates (Mean), empirical standard deviations (EMP.SE), mean estimated standard errors ($\widehat{SE}$) and empirical coverage rate of 95% confidence intervals (CP95%) for β.

$\beta^0[\exp(\beta^0)]$	$M^\star$	Method	Mean	EMP.SE	$\widehat{SE}$	CP95%
-1.946	2	LVCF	-1.658	0.249	0.248	0.761
[0.14]		MidI	-2.198	0.240	0.244	0.857
		PH-OC	-2.006	0.269	0.266	0.953
		PH-RSC	-2.003	0.268	0.266	0.952
	5	LVCF	-1.848	0.233	0.219	0.901
		MidI	-2.091	0.227	0.216	0.905
		PH-OC	-1.968	0.237	0.225	0.938
		PH-RSC	-1.969	0.237	0.225	0.938
	10	LVCF	-1.903	0.209	0.211	0.943
		MidI	-2.034	0.205	0.210	0.947
		PH-OC	-1.961	0.211	0.214	0.957
		PH-RSC	-1.962	0.211	0.214	0.957
-1.609	2	LVCF	-1.327	0.220	0.217	0.712
[0.20]		MidI	-1.867	0.206	0.213	0.800
		PH-OC	-1.638	0.241	0.238	0.953
		PH-RSC	-1.634	0.240	0.238	0.953
	5	LVCF	-1.521	0.194	0.195	0.908
		MidI	-1.771	0.188	0.192	0.897
		PH-OC	-1.632	0.200	0.202	0.950
		PH-RSC	-1.633	0.200	0.202	0.951
	10	LVCF	-1.571	0.192	0.187	0.944
		MidI	-1.706	0.188	0.185	0.935
		PH-OC	-1.625	0.195	0.190	0.949
		PH-RSC	-1.626	0.195	0.190	0.948
-0.693	2	LVCF	-0.534	0.162	0.161	0.824
[0.50]		MidI	-1.081	0.147	0.157	0.279
		PH-OC	-0.702	0.191	0.193	0.946
		PH-RSC	-0.701	0.190	0.192	0.949
	5	LVCF	-0.626	0.142	0.147	0.928
		MidI	-0.903	0.136	0.143	0.705
		PH-OC	-0.693	0.152	0.159	0.956
		PH-RSC	-0.693	0.152	0.158	0.956
	10	LVCF	-0.666	0.144	0.142	0.938
		MidI	-0.818	0.139	0.140	0.864
		PH-OC	-0.699	0.147	0.148	0.949
		PH-RSC	-0.700	0.147	0.147	0.948

β^0[exp(β^0)]	$M^\star$	Method	Mean	EMP.SE	$\widehat{SE}$	CP95%
0.000	2	LVCF	-0.002	0.141	0.135	0.944
[1.00]		MidI	-0.554	0.130	0.131	0.011
		PH-OC	0.003	0.183	0.178	0.950
		PH-RSC	0.004	0.183	0.177	0.952
	5	LVCF	0.003	0.122	0.124	0.955
		MidI	-0.300	0.113	0.121	0.270
		PH-OC	0.007	0.138	0.146	0.956
		PH-RSC	0.007	0.138	0.142	0.956
	10	LVCF	-0.004	0.121	0.121	0.960
		MidI	-0.177	0.116	0.119	0.686
		PH-OC	-0.003	0.129	0.130	0.954
		PH-RSC	-0.003	0.129	0.130	0.952
0.693	2	LVCF	0.462	0.119	0.118	0.498
[2.00]		MidI	-0.097	0.109	0.115	0.000
		PH-OC	0.680	0.175	0.179	0.936
		PH-RSC	0.684	0.177	0.175	0.938
	5	LVCF	0.572	0.107	0.110	0.810
		MidI	0.228	0.100	0.107	0.004
		PH-OC	0.690	0.132	0.145	0.958
		PH-RSC	0.689	0.132	0.137	0.957
	10	LVCF	0.627	0.104	0.107	0.910
		MidI	0.421	0.099	0.105	0.260
		PH-OC	0.694	0.118	0.126	0.962
		PH-RSC	0.693	0.117	0.122	0.962
1.609	2	LVCF	0.968	0.110	0.109	0.000
[5.00]		MidI	0.368	0.096	0.107	0.000
		PH-OC	1.472	0.179	0.210	0.869
		PH-RSC	1.516	0.190	0.195	0.897
	5	LVCF	1.212	0.096	0.099	0.013
		MidI	0.787	0.086	0.097	0.000
		PH-OC	1.577	0.139	0.166	0.951
		PH-RSC	1.575	0.137	0.151	0.948
	10	LVCF	1.365	0.100	0.097	0.277
		MidI	1.086	0.092	0.095	0.000
		PH-OC	1.598	0.127	0.131	0.949
		PH-RSC	1.594	0.125	0.126	0.940
1.946	2	LVCF	1.130	0.112	0.110	0.000
[7.00]		MidI	0.501	0.094	0.108	0.000
		PH-OC	1.695	0.182	0.230	0.723
		PH-RSC	1.773	0.198	0.205	0.816
	5	LVCF	1.410	0.097	0.098	0.000
		MidI	0.939	0.086	0.096	0.000
		PH-OC	1.890	0.148	0.187	0.933
		PH-RSC	1.890	0.147	0.158	0.929
	10	LVCF	1.600	0.095	0.095	0.050
		MidI	1.282	0.087	0.093	0.000
		PH-OC	1.927	0.127	0.143	0.946
		PH-RSC	1.920	0.124	0.130	0.934

Table A.3: Simulation study results when the distribution of V does not depend on covariates. Methods compared are LVCF, MidI, parametric Weibull calibration (WB-OC) and risk-set calibration models (WB-RSC), and nonparametric calibration (NP-OC) and risk-set calibration models (NP-RSC). The table presents mean estimates (Mean), empirical standard deviations (EMP.SE), mean estimated standard errors ($\widehat{SE}$) and empirical coverage rate of 95% confidence intervals (CP95%) for β.

			Weibull distribution for V				Piecewise exponential distribution for V			
β^0 [exp(β^0)]	$M^\star$	Method	Mean	EMP.SE	$\widehat{SE}$	CP95%	Mean	EMP.SE	$\widehat{SE}$	CP95%
-1.609	2	LVCF	-1.495	0.377	0.377	0.917	-1.487	0.407	1.454	0.916
[0.20]		MidI	-2.130	0.369	0.374	0.819	-2.203	0.396	1.744	0.753
		WB-OC	-1.659	0.392	0.392	0.972	-1.730	0.433	0.421	0.976
		WB-RSC	-1.662	0.393	0.393	0.970	-1.731	0.434	0.421	0.972
		NP-OC	-1.684	0.408	0.445	0.977	-1.708	0.438	0.462	0.973
		NP-RSC	1.671	0.405	0.441	0.972	-1.689	0.433	0.458	0.972
	5	LVCF	-1.572	0.304	0.306	0.946	-1.590	0.325	0.316	0.947
		MidI	-1.869	0.298	0.304	0.936	-1.899	0.316	0.313	0.916
		WB-OC	-1.638	0.309	0.312	0.963	-1.677	0.332	0.324	0.959
		WB-RSC	-1.643	0.310	0.312	0.964	-1.679	0.333	0.325	0.959
		NP-OC	-1.643	0.312	0.338	0.973	-1.656	0.329	0.346	0.970
		NP-RSC	-1.640	0.312	0.338	0.976	-1.653	0.328	0.345	0.972
	10	LVCF	-1.605	0.281	0.287	0.956	-1.612	0.311	0.296	0.946
		MidI	-1.759	0.277	0.286	0.965	-1.765	0.307	0.294	0.947
		WB-OC	-1.637	0.283	0.290	0.968	-1.651	0.315	0.299	0.950
		WB-RSC	-1.640	0.284	0.290	0.969	-1.652	0.315	0.299	0.948
		NP-OC	-1.637	0.284	0.308	0.975	-1.639	0.313	0.318	0.962
		NP-RSC	-1.636	0.284	0.308	0.971	-1.638	0.313	0.316	0.961
-0.693	2	LVCF	-0.618	0.255	0.254	0.935	-0.619	0.270	0.266	0.928
[0.50]		MidI	-1.255	0.242	0.250	0.365	-1.338	0.250	0.260	0.239
		WB-OC	-0.705	0.278	0.278	0.950	-0.745	0.304	0.307	0.950
		WB-RSC	-0.706	0.279	0.278	0.949	-0.745	0.304	0.307	0.950
		NP-OC	-0.710	0.285	0.299	0.961	-0.734	0.301	0.320	0.971
		NP-RSC	-0.703	0.282	0.297	0.955	-0.725	0.298	0.317	0.969
	5	LVCF	-0.665	0.206	0.211	0.951	-0.659	0.223	0.217	0.932
		MidI	-0.980	0.198	0.208	0.766	-0.992	0.209	0.213	0.734
		WB-OC	-0.704	0.213	0.220	0.964	-0.709	0.235	0.230	0.943
		WB-RSC	-0.706	0.214	0.220	0.964	-0.709	0.235	0.229	0.942
		NP-OC	-0.704	0.215	0.227	0.968	-0.698	0.231	0.233	0.947
		NP-RSC	-0.703	0.215	0.227	0.968	-0.696	0.231	0.233	0.945
	10	LVCF	-0.682	0.201	0.198	0.945	-0.676	0.196	0.203	0.968
		MidI	-0.848	0.196	0.196	0.901	-0.845	0.189	0.201	0.925
		WB-OC	-0.701	0.205	0.202	0.940	-0.699	0.201	0.209	0.964
		WB-RSC	-0.703	0.206	0.202	0.940	-0.699	0.201	0.209	0.964
		NP-OC	-0.701	0.206	0.206	0.945	-0.692	0.199	0.212	0.966
		NP-RSC	-0.699	0.206	0.207	0.948	-0.691	0.199	0.211	0.971

β^0 [exp(β^0)]	$M^\star$	Method	Weibull distribution for V				Piecewise exponential distribution for V			
			Mean	EMP.SE	$\widehat{SE}$	CP95%	Mean	EMP.SE	$\widehat{SE}$	CP95%
0.000	2	LVCF	-0.023	0.206	0.204	0.951	-0.008	0.211	0.212	0.953
[1.00]		MidI	-0.667	0.190	0.200	0.039	-0.733	0.192	0.206	0.022
		WB-OC	-0.020	0.235	0.234	0.956	0.003	0.253	0.261	0.951
		WB-RSC	-0.020	0.235	0.234	0.956	0.003	0.252	0.260	0.952
		NP-OC	-0.021	0.241	0.244	0.960	0.003	0.247	0.262	0.966
		NP-RSC	-0.019	0.242	0.244	0.960	0.006	0.247	0.262	0.956
	5	LVCF	-0.003	0.169	0.168	0.958	-0.003	0.170	0.174	0.952
		MidI	-0.341	0.158	0.166	0.472	-0.366	0.156	0.170	0.415
		WB-OC	-0.003	0.180	0.180	0.962	-0.003	0.185	0.191	0.961
		WB-RSC	-0.003	0.181	0.180	0.961	-0.003	0.184	0.190	0.960
		NP-OC	-0.001	0.181	0.183	0.953	0.000	0.181	0.190	0.955
		NP-RSC	-0.000	0.181	0.183	0.952	0.001	0.181	0.190	0.956
	10	LVCF	-0.006	0.162	0.158	0.952	0.005	0.167	0.163	0.944
		MidI	-0.190	0.155	0.156	0.798	-0.186	0.159	0.161	0.798
		WB-OC	-0.006	0.168	0.164	0.952	0.005	0.174	0.171	0.947
		WB-RSC	-0.006	0.168	0.164	0.952	0.005	0.173	0.170	0.946
		NP-OC	-0.004	0.168	0.166	0.953	0.007	0.172	0.171	0.952
		NP-RSC	-0.003	0.168	0.166	0.951	0.008	0.172	0.171	0.945
0.693	2	LVCF	0.567	0.178	0.175	0.906	0.551	0.182	0.182	0.881
[2.00]		MidI	-0.101	0.157	0.172	0.000	-0.188	0.156	0.177	0.000
		WB-OC	0.681	0.209	0.211	0.955	0.703	0.229	0.240	0.952
		WB-RSC	0.680	0.209	0.211	0.950	0.695	0.226	0.234	0.953
		NP-OC	0.674	0.207	0.215	0.958	0.679	0.220	0.229	0.956
		NP-RSC	0.684	0.211	0.219	0.959	0.692	0.229	0.237	0.952
	5	LVCF	0.630	0.142	0.143	0.929	0.627	0.151	0.149	0.925
		MidI	0.248	0.126	0.140	0.072	0.213	0.135	0.145	0.063
		WB-OC	0.689	0.157	0.160	0.956	0.701	0.172	0.170	0.943
		WB-RSC	0.686	0.156	0.159	0.956	0.691	0.168	0.167	0.946
		NP-OC	0.688	0.156	0.161	0.956	0.691	0.169	0.168	0.945
		NP-RSC	0.690	0.157	0.161	0.954	0.693	0.170	0.168	0.947
	10	LVCF	0.661	0.131	0.133	0.951	0.652	0.139	0.139	0.946
		MidI	0.445	0.122	0.132	0.540	0.425	0.128	0.137	0.500
		WB-OC	0.695	0.139	0.142	0.952	0.691	0.148	0.150	0.954
		WB-RSC	0.692	0.138	0.142	0.954	0.685	0.146	0.148	0.952
		NP-OC	0.696	0.139	0.143	0.955	0.689	0.147	0.149	0.949
		NP-RSC	0.697	0.140	0.143	0.951	0.690	0.147	0.150	0.955
1.609	2	LVCF	1.339	0.172	0.171	0.640	1.292	0.181	0.177	0.555
[5.00]		MidI	0.534	0.136	0.169	0.000	0.452	0.136	0.174	0.000
		WB-OC	1.537	0.206	0.206	0.931	1.549	0.230	0.223	0.935
		WB-RSC	1.548	0.213	0.210	0.932	1.529	0.226	0.222	0.936
		NP-OC	1.497	0.197	0.203	0.929	1.491	0.211	0.214	0.920
		NP-RSC	1.554	0.218	0.223	0.952	1.562	0.241	0.240	0.944
	5	LVCF	1.444	0.131	0.131	0.759	1.396	0.145	0.137	0.667
		MidI	0.932	0.103	0.130	0.000	0.852	0.116	0.134	0.000
		WB-OC	1.599	0.152	0.154	0.946	1.574	0.173	0.164	0.923
		WB-RSC	1.582	0.149	0.152	0.944	1.538	0.165	0.159	0.918
		NP-OC	1.578	0.148	0.153	0.945	1.562	0.169	0.163	0.926
		NP-RSC	1.591	0.152	0.157	0.949	1.574	0.172	0.167	0.940
	10	LVCF	1.502	0.120	0.119	0.852	1.477	0.131	0.125	0.805
		MidI	1.192	0.105	0.118	0.040	1.151	0.115	0.123	0.032
		WB-OC	1.605	0.136	0.133	0.946	1.591	0.147	0.141	0.938
		WB-RSC	1.589	0.132	0.131	0.943	1.561	0.141	0.137	0.929
		NP-OC	1.595	0.134	0.133	0.940	1.591	0.146	0.142	0.938
		NP-RSC	1.600	0.135	0.134	0.948	1.597	0.147	0.143	0.939

Table A.4: CMV data analysis results. We compare LVCF, MidI, NP-OC and NP-RSC. $N = 221$, 37 events for our data analysis

	Urine shedding				Blood shedding			
	Est	HR	$\widehat{\text{SE}}(Est)$	CI for HR	Est	HR	$\widehat{\text{SE}}(Est)$	CI for HR
LVCF	1.52	4.57	0.431	(1.96, 10.64)	2.64	13.95	0.336	(7.25, 27.07)
MidI	1.46	4.31	0.429	(1.86, 9.98)	2.44	11.51	0.333	(5.97, 22.04)
NP-OC	2.07	7.92	0.78	(1.41, 34.68)	2.27	9.69	0.493	(5.28, 39.18)
NP-RSC	2.15	8.61	0.75	(1.42, 29.03)	2.7	14.82	0.547	(6.5, 59.64)
Goggins et al.	1.95	7.03	0.384	(3.31, 14.92)	2.82	16.77	0.257	(11.02, 29.97)