BIOMETRICS , 1–12 DOI: 000
August 2014

Case-base methods for studying vaccination safety – Supplementary Materials

Olli Saarela[1,*], and James A. Hanley[2]

[1]Dalla Lana School of Public Health, University of Toronto,

155 College Street, 6th floor, Toronto, Ontario M5T 3M7, Canada.

[2]Department of Epidemiology, Biostatistics and Occupational Health, McGill University,

Purvis Hall, 1020 Pine Avenue West, Montreal, Quebec H3A 1A2, Canada.

[*]*email:* olli.saarela@utoronto.ca

This paper has been submitted for consideration for publication in *Biometrics*

Supplementary Appendix A: Adjusted variance estimators

Parametric two-step estimation

Using a parametric specification of the vaccination time distribution, in the continuous time setting one may use a two-step approach based on the conditional log-likelihood

$$L(\theta;\gamma,\alpha) \overset{\theta}{\propto} \prod_{i=1}^{n} \frac{[\theta\pi(t_i,\gamma,\alpha)]^{Z_i(t_i)}}{1-\pi(t_i,\gamma,\alpha)+\theta\pi(t_i,\gamma,\alpha)},$$

where n is the total number of cases and $Z_i(t_i)$ is the exposure status of the individual i at the event time t_i. In the following, since the normal approximation of $\log L(\theta;\gamma,\alpha)$ will work better on the $\eta = \log(\theta)$ scale, all the calculations will be carried out on the log-rate ratio scale. This may be estimated as $\hat{\eta}(\hat{\gamma},\hat{\alpha}) \equiv \arg\max_{\eta} L(\eta;\hat{\gamma},\hat{\alpha})$, where $(\hat{\gamma},\hat{\alpha})$ is the plug-in maximum likelihood estimator of (γ,α). Introduce now the following notations:

$$U^{\eta}(\eta;\gamma,\alpha) \equiv \partial \log L(\eta;\gamma,\alpha)/\partial\eta,$$

$$I^{\eta\eta}(\eta;\gamma,\alpha) \equiv \partial^2 \log L(\eta;\gamma,\alpha)/\partial\eta^2,$$

$$I^{\eta\gamma}(\eta;\gamma,\alpha) \equiv \partial^2 \log L(\eta;\gamma,\alpha)/\partial\eta\partial\gamma,$$

$$I^{\eta\alpha}(\eta;\gamma,\alpha) \equiv \partial^2 \log L(\eta;\gamma,\alpha)/\partial\eta\partial\alpha,$$

$$U^{\gamma}(\gamma) \equiv \partial \log L(\gamma)/\partial\gamma,$$

$$I^{\gamma\gamma}(\gamma) \equiv \partial^2 \log L(\gamma)/\partial\gamma^2$$

$$U^{\alpha}(\alpha) \equiv \partial \log L(\alpha)/\partial\alpha,$$

$$I^{\alpha\alpha}(\alpha) \equiv \partial^2 \log L(\alpha)/\partial\alpha^2.$$

In the following, we assume that the overlap of individuals between the sampled risksets is negligible, so that they can be assumed independent, and that the parametric models are correctly specified. (The first assumption is reasonable in the vaccination safety context where the size of the study population N is usually much larger than the number of cases

n.) Then

$$E[U^\gamma(\gamma)] = \sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_{v_{ij}}\sum_{a_{ij}=0}^{1} a_{ij}\frac{\partial P(\mathrm{d}v_{ij} \mid A_{ij}=1,\gamma)/\partial\gamma}{P(\mathrm{d}v_{ij} \mid A_{ij}=1,\gamma)}P(\mathrm{d}v_{ij} \mid a_{ij})P(a_{ij})$$
$$= \sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_{v_{ij}}\frac{\partial P(\mathrm{d}v_{ij} \mid A_{ij}=1,\gamma)/\partial\gamma}{P(\mathrm{d}v_{ij} \mid A_{ij}=1,\gamma)}P(\mathrm{d}v_{ij} \mid A_{ij}=1)P(A_{ij}=1) = 0,$$

$E[U_{ij}^\gamma(\gamma) \mid A_{ij}=0]=0$, $E[U_{ij}^\gamma(\gamma) \mid A_{ij}=1]=0$ and

$$E[U^\gamma(\gamma)U^\alpha(\alpha)'] = \sum_{i=1}^{n}\sum_{j=1}^{m_i} E\left[U_{ij}^\gamma(\gamma)U_{ij}^\alpha(\alpha)'\right]$$
$$= \sum_{i=1}^{n}\sum_{j=1}^{m_i} E_{A_{ij}}\left[E\left\{U_{ij}^\gamma(\gamma) \mid A_{ij}\right\}U_{ij}^\alpha(\alpha)'\right] = 0.$$

We are interested in the asymptotic variance of $\hat\eta(\hat\gamma,\hat\alpha)$. A first order Taylor expansion around the true parameter values $(\eta_0,\gamma_0,\alpha_0)$ gives

$$\begin{aligned}
0 &= \frac{1}{n}U^\eta(\hat\eta;\hat\gamma,\hat\alpha)\\
&\approx \frac{1}{n}U^\eta(\eta_0;\gamma_0,\alpha_0) + \frac{1}{n}I^{\eta\eta}(\eta_0;\gamma_0,\alpha_0)(\hat\eta-\eta_0) + \frac{1}{n}I^{\eta\gamma}(\eta_0;\gamma_0,\alpha_0)(\hat\gamma-\gamma_0)\\
&\quad + \frac{1}{n}I^{\eta\alpha}(\eta_0;\gamma_0,\alpha_0)(\hat\alpha-\alpha_0)\\
&\approx \frac{1}{n}U^\eta(\eta_0;\gamma_0,\alpha_0) + E[I_i^{\eta\eta}(\eta_0;\gamma_0,\alpha_0)](\hat\eta-\eta_0) + E[I_i^{\eta\gamma}(\eta_0;\gamma_0,\alpha_0)](\hat\gamma-\gamma_0)\\
&\quad + E[I_i^{\eta\alpha}(\eta_0;\gamma_0,\alpha_0)](\hat\alpha-\alpha_0)
\end{aligned}$$

or

$$\sqrt{n}(\hat\eta-\eta_0) \approx E[-I_i^{\eta\eta}(\eta_0;\gamma_0,\alpha_0)]^{-1}\Bigg[\frac{\sqrt{n}}{n}U^\eta(\eta_0;\gamma_0,\alpha_0) + E[I_i^{\eta\gamma}(\eta_0;\gamma_0,\alpha_0)]\sqrt{n}(\hat\gamma-\gamma_0)$$
$$+ E[I_i^{\eta\alpha}(\eta_0;\gamma_0,\alpha_0)]\sqrt{n}(\hat\alpha-\alpha_0)\Bigg].$$

Here, by further first order expansions around γ_0 and α_0,

$$\sqrt{n}(\hat\gamma-\gamma_0) \approx E[-I_i^{\gamma\gamma}(\gamma_0)]^{-1}\frac{\sqrt{n}}{n}U^\gamma(\gamma_0),$$

and

$$\sqrt{n}(\hat\alpha-\alpha_0) \approx E[-I_i^{\alpha\alpha}(\alpha_0)]^{-1}\frac{\sqrt{n}}{n}U^\alpha(\alpha_0),$$

which substituted back to the previous expression give

$$\sqrt{n}(\hat{\eta}-\eta_0) \approx E[-I_i^{\eta\eta}(\eta_0;\gamma_0,\alpha_0)]^{-1}\Bigg[\frac{\sqrt{n}}{n}\sum_{i=1}^{n}\Big\{U_i^{\eta}(\eta_0;\gamma_0,\alpha_0) + E[I_i^{\eta\gamma}(\eta_0;\gamma_0,\alpha_0)]E[-I_i^{\gamma\gamma}(\gamma_0)]^{-1}U_i^{\gamma}(\gamma_0) + E[I_i^{\eta\alpha}(\eta_0;\alpha_0,\alpha_0)]E[-I_i^{\alpha\alpha}(\alpha_0)]^{-1}U_i^{\alpha}(\alpha_0)\Big\}\Bigg].$$

By the central limit theorem it then follows that

$$\sqrt{n}(\hat{\eta}-\eta_0) \xrightarrow{\mathrm{d}} N(0, E[-I_i^{\eta\eta}(\eta_0;\gamma_0,\alpha_0)]^{-1}V[B_i(\eta_0,\gamma_0,\alpha_0)]E[-I^{\eta\eta}(\eta_0;\gamma_0,\alpha_0)]^{-1}),$$

where

$$B_i(\eta_0,\gamma_0,\alpha_0) \equiv U_i^{\eta}(\eta_0;\gamma_0,\alpha_0) + E[I_i^{\eta\gamma}(\eta_0;\gamma_0,\alpha_0)]E[-I_i^{\gamma\gamma}(\gamma_0)]^{-1}U_i^{\gamma}(\gamma_0) + E[I_i^{\eta\alpha}(\eta_0;\gamma_0,\alpha_0)]E[-I_i^{\alpha\alpha}(\alpha_0)]^{-1}U_i^{\alpha}(\alpha_0).$$

Since $E[B_i(\eta_0,\gamma_0,\alpha_0)] = 0$, the variance $V[B_i(\eta_0,\gamma_0,\alpha_0)]$ can be further written as

$$\begin{aligned}
&V[B_i(\eta_0,\gamma_0,\alpha_0)]\\
&= E[B_i(\eta_0,\gamma_0,\alpha_0)B_i(\eta_0,\gamma_0,\alpha_0)']\\
&= E[U_i^{\eta}(\eta_0;\gamma_0,\alpha_0)U_i^{\eta}(\eta_0;\gamma_0,\alpha_0)']\\
&\quad + E[U_i^{\eta}(\eta_0;\gamma_0,\alpha_0)U_i^{\gamma}(\gamma_0)']E[-I_i^{\gamma\gamma}(\gamma_0)]^{-1'}E[I_i^{\eta\gamma}(\eta_0;\gamma_0,\alpha_0)']\\
&\quad + E[U_i^{\eta}(\eta_0;\gamma_0,\alpha_0)U_i^{\alpha}(\alpha_0)']E[-I_i^{\alpha\alpha}(\alpha_0)]^{-1'}E[I_i^{\eta\alpha}(\eta_0;\gamma_0,\alpha_0)']\\
&\quad + E[I_i^{\eta\gamma}(\eta_0;\gamma_0,\alpha_0)]E[-I_i^{\gamma\gamma}(\gamma_0)]^{-1}E[U_i^{\gamma}(\gamma_0)U_i^{\eta}(\eta_0;\gamma_0,\alpha_0)']\\
&\quad + E[I_i^{\eta\gamma}(\eta_0;\gamma_0,\alpha_0)]E[-I_i^{\gamma\gamma}(\gamma_0)]^{-1}E[U_i^{\gamma}(\gamma_0)U_i^{\gamma}(\gamma_0)']E[-I_i^{\gamma\gamma}(\gamma_0)]^{-1}E[I_i^{\eta\gamma}(\eta_0;\gamma_0,\alpha_0)']\\
&\quad + E[I_i^{\eta\gamma}(\eta_0;\gamma_0,\alpha_0)]E[-I_i^{\gamma\gamma}(\gamma_0)]^{-1}E[U_i^{\gamma}(\gamma_0)U_i^{\alpha}(\alpha_0)']E[-I_i^{\alpha\alpha}(\alpha_0)]^{-1}E[I_i^{\eta\alpha}(\eta_0;\alpha_0,\alpha_0)']\\
&\quad + E[I_i^{\eta\alpha}(\eta_0;\gamma_0,\alpha_0)]E[-I_i^{\alpha\alpha}(\alpha_0)]^{-1}E[U_i^{\alpha}(\alpha_0)U_i^{\eta}(\eta_0;\gamma_0,\alpha_0)']\\
&\quad + E[I_i^{\eta\alpha}(\eta_0;\gamma_0,\alpha_0)]E[-I_i^{\alpha\alpha}(\alpha_0)]^{-1}E[U_i^{\alpha}(\alpha_0)U_i^{\gamma}(\gamma_0)']E[-I_i^{\gamma\gamma}(\gamma_0)]^{-1}E[I_i^{\eta\gamma}(\eta_0;\gamma_0,\alpha_0)']\\
&\quad + E[I_i^{\eta\alpha}(\eta_0;\gamma_0,\alpha_0)]E[-I_i^{\alpha\alpha}(\alpha_0)]^{-1}E[U_i^{\alpha}(\alpha_0)U_i^{\alpha}(\alpha_0)']E[-I_i^{\alpha\alpha}(\alpha_0)]^{-1}E[I_i^{\eta\alpha}(\eta_0;\alpha_0,\alpha_0)'],
\end{aligned}$$

where the fifth term

$$E[I_i^{\eta\gamma}(\eta_0;\gamma_0,\alpha_0)]E[-I_i^{\gamma\gamma}(\gamma_0)]^{-1}E[U_i^{\gamma}(\gamma_0)U_i^{\gamma}(\gamma_0)']E[-I_i^{\gamma\gamma}(\gamma_0)]^{-1}E[I_i^{\eta\gamma}(\eta_0;\gamma_0,\alpha_0)']$$
$$= E[I_i^{\eta\gamma}(\eta_0;\gamma_0,\alpha_0)]E[-I_i^{\gamma\gamma}(\gamma_0)]^{-1}E[I_i^{\eta\gamma}(\eta_0;\gamma_0,\alpha_0)'],$$

since $\log L(\gamma)$ is a log-likelihood for which $E[U_i^{\gamma}(\gamma_0)U_i^{\gamma}(\gamma_0)'] = E[-I_i^{\gamma\gamma}(\gamma_0)]$. Similarly for the final term,

$$E[I_i^{\eta\alpha}(\eta_0;\gamma_0,\alpha_0)]E[-I_i^{\alpha\alpha}(\alpha_0)]^{-1}E[U_i^{\alpha}(\alpha_0)U_i^{\alpha}(\alpha_0)']E[-I_i^{\alpha\alpha}(\alpha_0)]^{-1}E[I_i^{\eta\alpha}(\eta_0;\gamma_0,\alpha_0)']$$
$$= E[I_i^{\eta\alpha}(\eta_0;\gamma_0,\alpha_0)]E[-I_i^{\alpha\alpha}(\alpha_0)]^{-1}E[I_i^{\eta\alpha}(\eta_0;\gamma_0,\alpha_0)'].$$

If we further assume that the exposure and outcome status of case i are independent of the vaccination times of the m_i time-matched controls (a standard IID/exchangeability assumption), we have further that $U_i^{\eta}(\eta;\gamma,\alpha) \perp\!\!\!\perp (U_i^{\gamma}(\gamma), U_i^{\alpha}(\alpha))$, $E[U_i^{\eta}(\eta;\gamma,\alpha)U_i^{\gamma}(\gamma)'] = E[U_i^{\eta}(\eta;\gamma,\alpha)]E[U_i^{\gamma}(\gamma)'] = 0$ and $E[U_i^{\eta}(\eta;\gamma,\alpha)U_i^{\alpha}(\alpha)'] = 0$. Thus, the variance simplifies into

$$\begin{aligned} V[B_i(\eta_0,\gamma_0,\alpha_0)] = {} & E[U_i^{\eta}(\eta_0;\gamma_0,\alpha_0)U_i^{\eta}(\eta_0;\gamma_0,\alpha_0)'] \\ & + E[I_i^{\eta\gamma}(\eta_0;\gamma_0,\alpha_0)]E[-I_i^{\gamma\gamma}(\gamma_0)]^{-1}E[I_i^{\eta\gamma}(\eta_0;\gamma_0,\alpha_0)'] \\ & + E[I_i^{\eta\alpha}(\eta_0;\gamma_0,\alpha_0)]E[-I_i^{\alpha\alpha}(\alpha_0)]^{-1}E[I_i^{\eta\alpha}(\eta_0;\gamma_0,\alpha_0)']. \end{aligned}$$

Finally, since the information equality $E[U_i^{\eta}(\eta_0;\gamma_0,\alpha_0)U_i^{\eta}(\eta_0;\gamma_0,\alpha_0)'] = E[-I_i^{\eta\eta}(\eta_0;\gamma_0,\alpha_0)]$ applies also for the conditional likelihood $L(\eta;\gamma,\alpha)$ we have motivated the variance estimator

$$\begin{aligned} \hat{V}[\hat{\eta}] = {} & -I^{\eta\eta}(\hat{\eta};\hat{\gamma},\hat{\alpha})^{-1} + I^{\eta\eta}(\hat{\eta};\hat{\gamma},\hat{\alpha})^{-1}I^{\eta\gamma}(\hat{\eta};\hat{\gamma},\hat{\alpha})[-I^{\gamma\gamma}(\hat{\gamma})]^{-1}I^{\eta\gamma}(\hat{\eta};\hat{\gamma},\hat{\alpha})'I^{\eta\eta}(\hat{\eta};\hat{\gamma},\hat{\alpha})^{-1} \\ & + I^{\eta\eta}(\hat{\eta};\hat{\gamma},\hat{\alpha})^{-1}I^{\eta\alpha}(\hat{\eta};\hat{\gamma},\hat{\alpha})[-I^{\alpha\alpha}(\hat{\alpha})]^{-1}I^{\eta\alpha}(\hat{\eta};\hat{\gamma},\hat{\alpha})'I^{\eta\eta}(\hat{\eta};\hat{\gamma},\hat{\alpha})^{-1}. \end{aligned} \tag{1}$$

Semiparametric two-step estimation

Moving onto a semi-parametric approach to the estimation problem, we first note that since $\pi(t) = P(A = 1 \mid \alpha)\int_{v\in(t-7,t]} P(V \in \mathrm{d}v \mid A = 1) = \alpha F(t) - \alpha F(t-7) \equiv \pi(t, F, \alpha)$, the exposure prevalence function is a deterministic function of the distribution function F for vaccination time and the eventual vaccination prevalence α. Denoting the total number of controls as $M \equiv \sum_{i=1}^{n} m_i$, the latter may be estimated by $\hat{\alpha} = \frac{1}{M}\sum_{i=1}^{n}\sum_{j=1}^{m_i} A_{ij}$ and the the distribution function in turn by $\hat{F}(t) \equiv \frac{1}{\hat{\alpha}M}\sum_{i=1}^{n}\sum_{j=1}^{m_i} A_{ij}\mathbf{1}_{\{V_{ij}<t\}}$. Analogously to above we take

$$
\begin{aligned}
U^{\eta}(\eta; F, \alpha) &\equiv \partial \log L(\eta; F, \alpha)/\partial\eta, \\
I^{\eta\eta}(\eta; F, \alpha) &\equiv \partial^2 \log L(\eta; F, \alpha)/\partial\eta^2, \\
I^{\eta F}(\eta; F, \alpha) &\equiv \partial^2 \log L(\eta; F, \alpha)/\partial\eta\partial F,
\end{aligned}
$$

where the last gradient vector is to be understood as pointwise differentiation of $U^{\eta}(\eta; F, \alpha)$ with respect to $F(t)$ at finitely many points t. Similarly as before we may now take $\hat{\eta}(\hat{F}, \hat{\alpha}) \equiv \arg\max_{\eta} L(\eta; \hat{F}, \hat{\alpha})$ and expand $U^{\eta}(\hat{\eta}; \hat{F}, \hat{\alpha})$ around the true values (η_0, F, α_0) as

$$
\begin{aligned}
0 &= \frac{1}{n}U^{\eta}(\hat{\eta}; \hat{F}, \hat{\alpha}) \\
&\approx \frac{1}{n}U^{\eta}(\eta_0; F, \alpha_0) + \frac{1}{n}I^{\eta\eta}(\eta_0; F, \alpha_0)(\hat{\eta} - \eta_0) + \frac{1}{n}I^{\eta F}(\eta_0; F, \alpha_0)(\hat{F} - F) \\
&\quad + \frac{1}{n}I^{\eta\alpha}(\eta_0; F, \alpha_0)(\hat{\alpha} - \alpha_0) \\
&\approx \frac{1}{n}U^{\eta}(\eta_0; F, \alpha_0) + E[I_i^{\eta\eta}(\eta_0; F, \alpha_0)](\hat{\eta} - \eta_0) + E[I_i^{\eta F}(\eta_0; F, \alpha_0)](\hat{F} - F) \\
&\quad + E[I_i^{\eta\alpha}(\eta_0; F, \alpha_0)](\hat{\alpha} - \alpha_0)
\end{aligned}
$$

or

$$
\begin{aligned}
\hat{\eta} - \eta_0 \approx E[-I_i^{\eta\eta}(\eta_0; F, \alpha_0)]^{-1}[\frac{1}{n}U^{\eta}(\eta_0; F, \alpha_0) &+ E[I_i^{\eta F}(\eta_0; F, \alpha_0)](\hat{F} - F) \\
&+ E[I_i^{\eta\alpha}(\eta_0; F, \alpha_0)](\hat{\alpha} - \alpha_0)]. \qquad (2)
\end{aligned}
$$

We know that here $\sqrt{\hat{\alpha}M}(\hat{F}(t) - F(t)) \xrightarrow{\mathrm{d}} N(0, F(t)(1 - F(t)))$ at any given point t (e.g. van der Vaart, 1998, p. 265). In addition, we can again assume that $U^{\eta}(\eta; F) \perp\!\!\!\perp (\hat{F}, \hat{\alpha})$. Further, we note that

$$
\begin{aligned}
&E\left[(\hat{F}(t)-F(t))(\hat{\alpha}-\alpha_0)\right]\\
&=E\left[\frac{\sum_{i=1}^{n}\sum_{j=1}^{m_i}(A_{ij}\mathbf{1}_{\{V_{ij}<t\}}-A_{ij}F(t))}{\sum_{i=1}^{n}\sum_{j=1}^{m_i}A_{ij}}(\hat{\alpha}-\alpha_0)\right]\\
&=E_A\left[E\left\{\frac{\sum_{i=1}^{n}\sum_{j=1}^{m_i}(A_{ij}\mathbf{1}_{\{V_{ij}<t\}}-A_{ij}F(t))}{\sum_{i=1}^{n}\sum_{j=1}^{m_i}A_{ij}}\mid A\right\}(\hat{\alpha}-\alpha_0)\right]\\
&=0.
\end{aligned}
$$

Hence, by noting that the right hand side of (2) has zero expectation, as in the previous section we obtain the variance approximation

$$
\begin{aligned}
V[\hat{\eta}] \approx\ & E[-I_i^{\eta\eta}(\eta_0;F,\alpha_0)]^{-1}\frac{1}{n^2}V[U^{\eta}(\eta_0;F,\alpha_0)]E[-I_i^{\eta\eta}(\eta_0;F,\alpha_0)']^{-1}\\
&+E[-I_i^{\eta\eta}(\eta_0;F,\alpha_0)]^{-1}E[I_i^{\eta F}(\eta_0;F,\alpha_0)]V[\hat{F}]E[I_i^{\eta F}(\eta_0;F,\alpha_0)']E[-I_i^{\eta\eta}(\eta_0;F,\alpha_0)']^{-1}\\
&+E[-I_i^{\eta\eta}(\eta_0;F,\alpha_0)]^{-1}E[I_i^{\eta\alpha}(\eta_0;F,\alpha_0)]V[\hat{\alpha}]E[I_i^{\eta\alpha}(\eta_0;F,\alpha_0)']E[-I_i^{\eta\eta}(\eta_0;F,\alpha_0)']^{-1},
\end{aligned}
$$

motivating the estimator

$$
\begin{aligned}
\hat{V}[\hat{\eta}] = & -I^{\eta\eta}(\hat{\eta};\hat{F},\hat{\alpha})^{-1}+I^{\eta\eta}(\hat{\eta};\hat{F},\hat{\alpha})^{-1}I^{\eta F}(\hat{\eta};\hat{F},\hat{\alpha})V[\hat{F}]I^{\eta F}(\hat{\eta};\hat{F},\hat{\alpha})'I^{\eta\eta}(\hat{\eta};\hat{F})^{-1}\\
&+I^{\eta\eta}(\hat{\eta};\hat{F},\hat{\alpha})^{-1}I^{\eta\alpha}(\hat{\eta};\hat{F},\hat{\alpha})V[\hat{\alpha}]I^{\eta\alpha}(\hat{\eta};\hat{F},\hat{\alpha})'I^{\eta\eta}(\hat{\eta};\hat{F},\hat{\alpha})^{-1}.
\end{aligned}
$$

Here the covariance terms in the variance-covariance matrix $V[\hat{F}]$ are given by $C(\hat{F}(t_i),\hat{F}(t_j)) = \frac{1}{\hat{\alpha}M}[F(t_i\wedge t_j)-F(t_i)F(t_j)]$ (van der Vaart, 1998, p. 266), and would in practice be estimated by substituting in the estimated $\hat{F}$ values. In addition, $V[\hat{\alpha}]\approx\hat{\alpha}(1-\hat{\alpha})/M$. The variance estimator obtained in the semi-parametric case is the direct analogy of (1).

The partial derivatives of

$$
\log L(\eta;F,\alpha)=\sum_{i=1}^{n}Z_i(t_i)\eta+\sum_{i=1}^{n}Z_i(t_i)\log[\pi(t_i,F,\alpha)]-\sum_{i=1}^{n}\log[1+(\exp(\eta)-1)\pi(t_i,F,\alpha)],
$$

w.r.t. η are given by

$$
U^{\eta}(\eta;F,\alpha)=\sum_{i=1}^{n}Z_i(t_i)-\sum_{i=1}^{n}\frac{\exp(\eta)\pi(t_i,F,\alpha)}{1+(\exp(\eta)-1)\pi(t_i,F,\alpha)},
$$

and

$$I^{\eta\eta}(\eta; F, \alpha) = \sum_{i=1}^{n} \frac{\exp(\eta)[\pi(t_i, F, \alpha)^2 - \pi(t_i, F, \alpha)]}{[1 + (\exp(\eta) - 1)\pi(t_i, F, \alpha)]^2}.$$

Finally, with

$$U^{\eta}(\eta; F, \alpha) = \sum_{i=1}^{n} Z_i(t_i) - \sum_{i=1}^{n} \frac{\exp(\eta)[\alpha F(t_i) - \alpha F(t_i - 7)]}{1 + (\exp(\eta) - 1)[\alpha F(t_i) - \alpha F(t_i - 7)]},$$

we have that

$$I^{\eta F(t_i)}(\eta; F, \alpha) = -\frac{\exp(\eta)\alpha}{[1 + (\exp(\eta) - 1)(\alpha F(t_i) - \alpha F(t_i - 7))]^2},$$

$$I^{\eta F(t_i - 7)}(\eta; F, \alpha) = \frac{\exp(\eta)\alpha}{[1 + (\exp(\eta) - 1)(\alpha F(t_i) - \alpha F(t_i - 7))]^2}$$

and

$$I^{\eta\alpha}(\eta; F, \alpha) = -\sum_{i=1}^{n} \frac{\exp(\eta)(F(t_i) - F(t_i - 7))}{[1 + (\exp(\eta) - 1)(\alpha F(t_i) - \alpha F(t_i - 7))]^2}.$$

Supplementary Appendix B: Conditional likelihood for self-matched case-base sampling

Let $Q_i(t) = N_i(t) + R_i(t)$ be a counting process for the total number of selected case and base series person-moments, and let $\mathrm{d}Q_i(t_{ij}) = 1 \; \forall \; j \in J$ indicate that all numbered person-moments $J = \{1, \ldots, m_i + 1\}$ at times $t_{i1}, \ldots, t_{i(m_i+1)}$ were selected in the case-base sample. It is assumed that the events are generated by a non-homogeneous Poisson process with rate $\lambda_{Z_i(t)}(t, \alpha_i) = P(\mathrm{d}N_i(t) \mid N_i(t-), Z_i(t))/\mathrm{d}t = \exp\{\alpha_i + f(t, \beta) + \eta Z_i(t)\}$. Since the likelihood construction involves conditioning on future exposure status, it is further required that $\mathrm{d}N_i(t) \perp\!\!\!\perp \{Z_i(u) : u \neq t\} \mid (N_i(t-), Z_i(t))$ (the incidence rate depends only on the current exposure). On the sampling mechanism we require that $\mathrm{d}R_i(t) \perp\!\!\!\perp \{N_i(u), Z_i(u) : 0 \leqslant u \leqslant \tau\}$, so that sampling of the base series is independent of the event and exposure history of individual i (cf. Section 2.2.2).

With the base series sampled with uniform probabilities (illustrated in the schematic in Figure a), and the proportional hazards model, we may now write a conditional logistic form

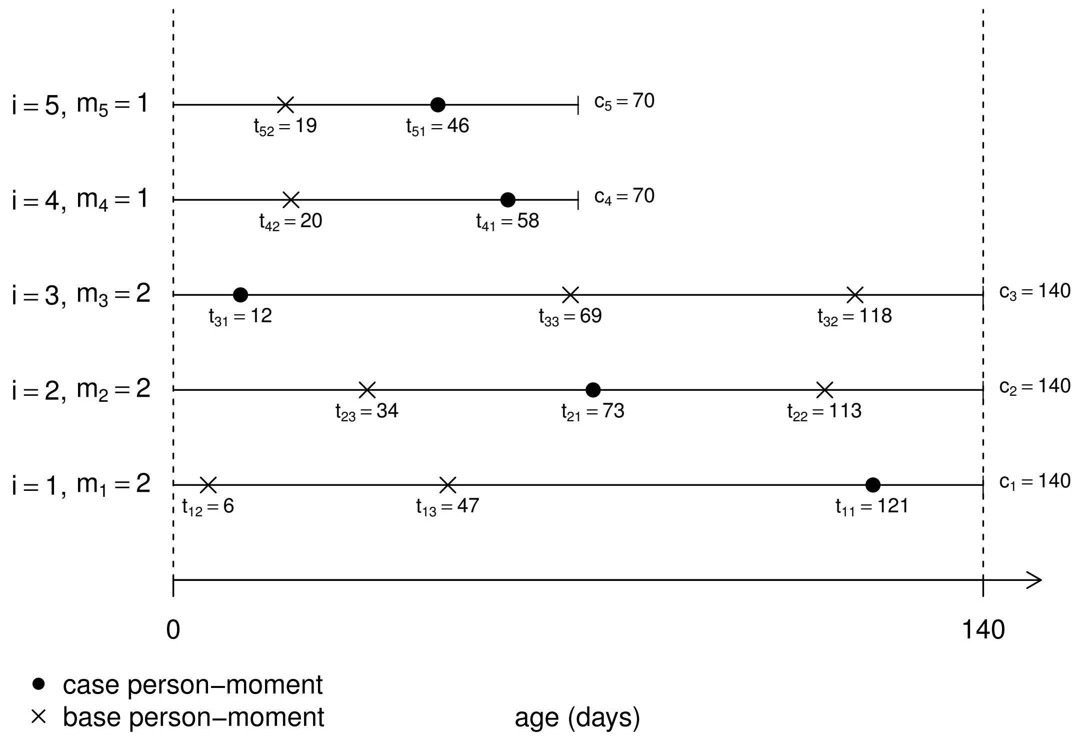

Figure a. A schematic of self-matched case-base sampling of person-moments. A base series of an expected size of $M = \sum_{i=1}^{n} E[m_i] = 8$ is drawn uniformly from a total of 560 days of follow-up contributed by $n = 5$ individuals with an outcome event.

likelihood contribution for the $m_i + 1$ person-moments contributed by an individual i with a single outcome event as

$$
\begin{aligned}
&P\left(\mathrm{d}N_i(t_{i1}) = 1 \mid \sum_{j=1}^{m_i+1} \mathrm{d}N_i(t_{ij}) = 1, Z_i(t_{i1}), \ldots, Z_i(t_{i(m_i+1)}), \mathrm{d}Q_i(t_{ij}) = 1 \; \forall \; j \in J\right) \\
&= \frac{P(\mathrm{d}N_i(t_{i1}) = 1, \mathrm{d}Q_i(t_{ij}) = 1 \; \forall \; j \in J \mid Z_i(t_{i1}), \ldots, Z_i(t_{i(m_i+1)}))}{\sum_{j=1}^{m_i+1} P(\mathrm{d}N_i(t_{ij}) = 1, \mathrm{d}Q_i(t_{il}) = 1 \; \forall \; l \in J \mid Z_i(t_{i1}), \ldots, Z_i(t_{i(m_i+1)}))} \\
&= \frac{P(\mathrm{d}R_i(t_{ij}) = 1 \; \forall \; j \in J \setminus \{1\}) P(\mathrm{d}N_i(t_{i1}) = 1 \mid N_i(t_{i1}-), Z_i(t_{i1}))}{\sum_{j=1}^{m_i+1} P(\mathrm{d}R_i(t_{il}) = 1 \; \forall \; l \in J \setminus \{j\}) P(\mathrm{d}N_i(t_{ij}) = 1 \mid N_i(t_{ij}-), Z_i(t_{ij}))} \\
&= \frac{\lambda_{Z_i(t_{i1})}(t_{i1}, \alpha_i)}{\sum_{j=1}^{m_i+1} \lambda_{Z_i(t_{ij})}(t_{ij}, \alpha_i)} \\
&= \frac{\exp\{f(t_{i1}, \beta) + \eta Z_i(t_{i1})\}}{\sum_{j=1}^{m_i+1} \exp\{f(t_{ij}, \beta) + \eta Z_i(t_{ij})\}}.
\end{aligned}
$$

For individuals with more than one event, the likelihood expression generalizes in the usual way as under group-matched case-control settings.

Supplementary Appendix C: Population-time presentation of time-matched and self-matched sampling of base series person-moments

To illustrate visually the concepts introduced in Section 1.2, here we present a step-by-step construction of the top panel of Figure 1, as well as time-matched and self-matched sampling of the base series. Figure b shows the study base, while Figure c splits this into exposed and unexposed population time. (It should be noted that this is only possible if the vaccination histories are known for the whole study population; however, an estimated version can be constructed using our proposed estimation approach as in Figure 2.) Figure d shows the case series, while Figure e adds in a one-to-one time-matched base series (one person-moment sampled from the risk set at each event time). Figure f presents the population-time contributed only by the individuals with an outcome event, and Figure g shows a self-matched base series drawn from this population-time.

References

van der Vaart, A. W. (1998). *Asymptotic statistics.* Cambridge University Press, Cambridge.

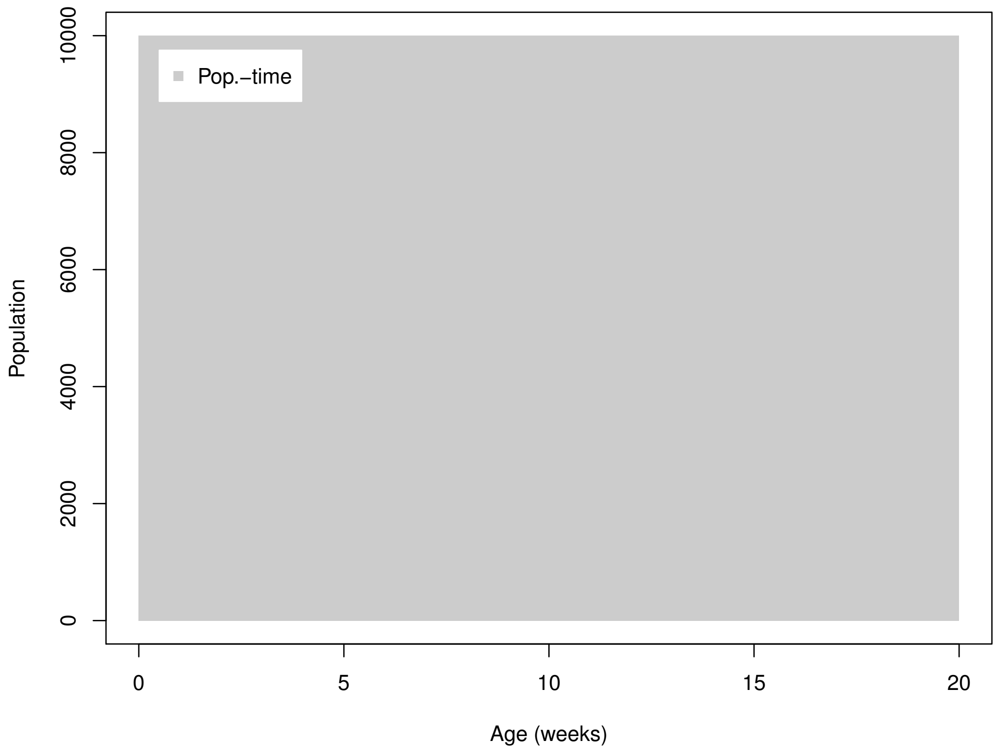

Figure b. Study base (200000 person-weeks of population-time).

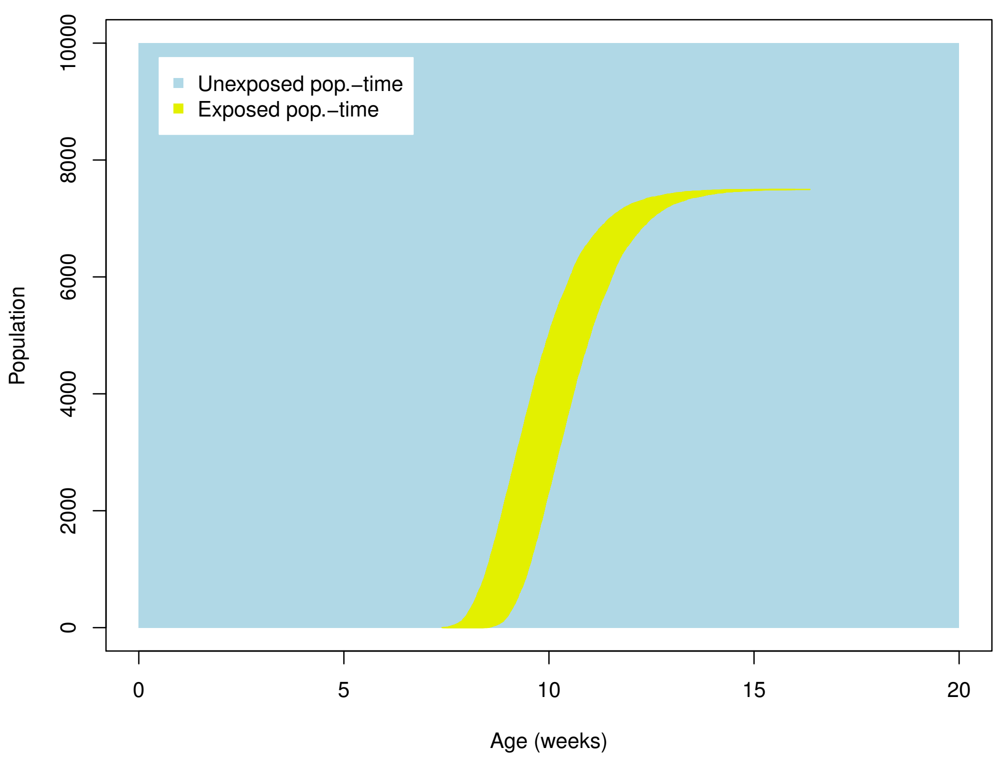

Figure c. Study base split into exposed and unexposed population-time.

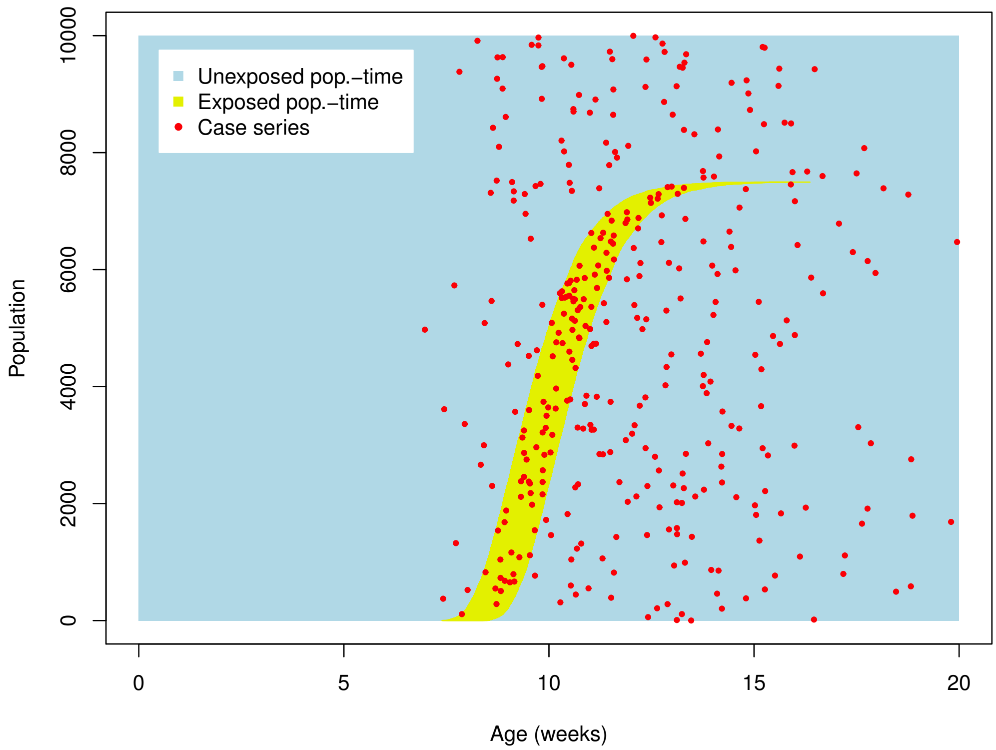

Figure d. Case series ($n = 374$).

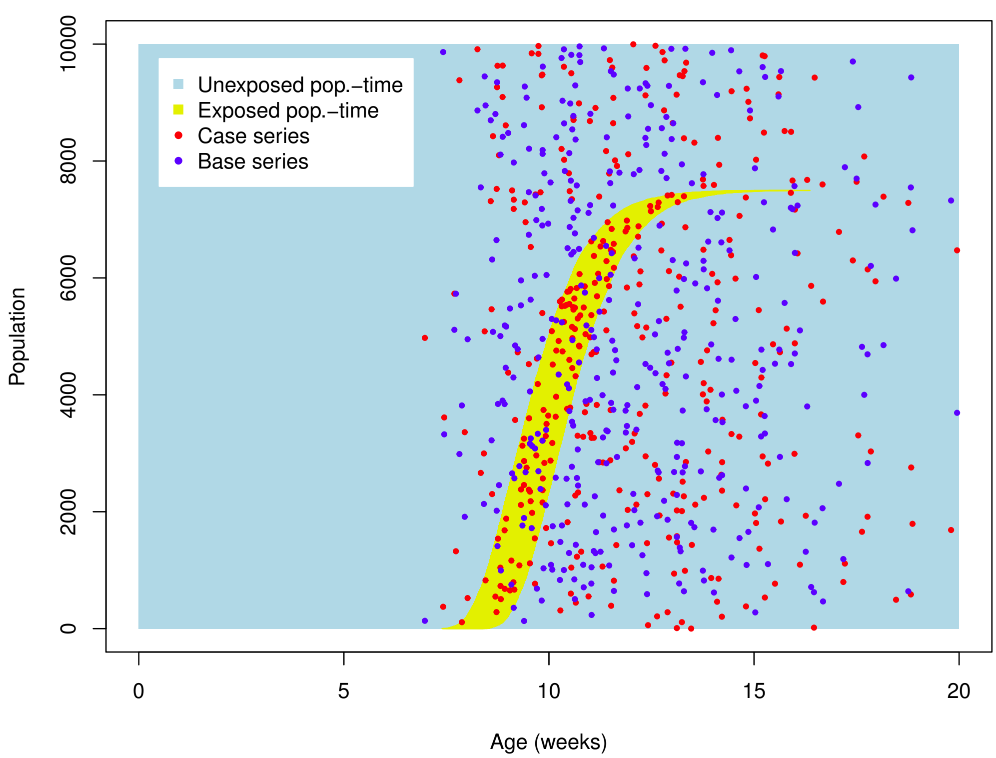

Figure e. One-to-one time-matched base series.

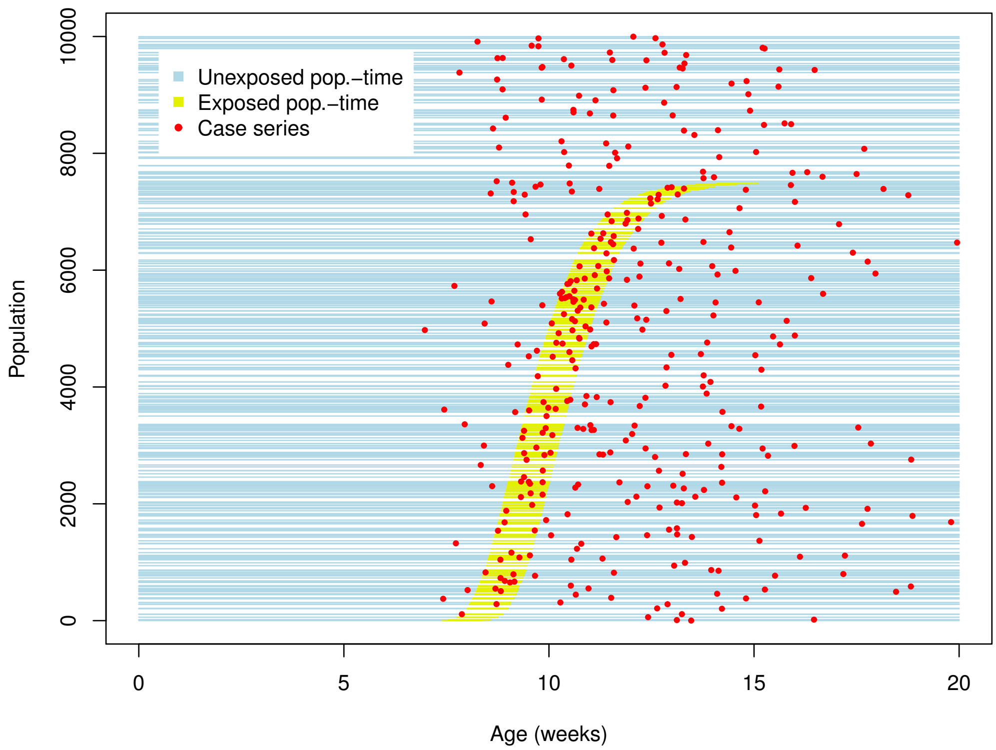

Figure f. Self-matched population-time.

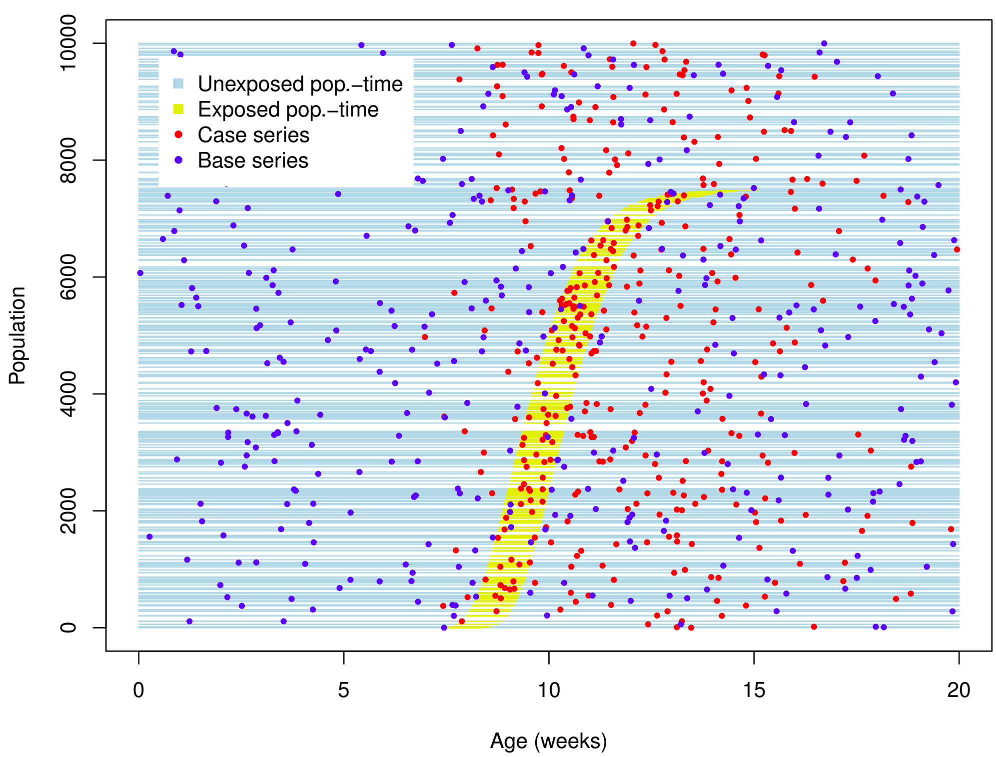

Figure g. One-to-one self-matched base series.