Biometrika (2012), **99**, 3, *pp.* 1–10

Printed in Great Britain

Supplementary material for "Nonparametric incidence estimation from prevalent cohort survival data"

By MARCO CARONE

Division of Biostatistics, School of Public Health, University of California Berkeley, 109 Haviland Hall, Berkeley, California 94720, U.S.A.

mcarone@berkeley.edu

MASOUD ASGHARIAN

Department of Mathematics and Statistics, McGill University, 805 Sherbrooke Street West, Montreal, Quebec H3A 2K6, Canada

masoud@math.mcgill.ca

and MEI-CHENG WANG

Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, 615 North Wolfe Street, Baltimore, Maryland 21205, U.S.A.

mcwang@jhsph.edu

Supplementary material

Additional technical details

Uniform consistency of $\hat{G}$

Lemma 4.2 of Wang (1991) establishes the pointwise consistency of $\hat{G}$. However, we require uniform consistency. We begin by proving the uniform consistency of $\hat{G}$ on $[0, a]$ for any $0 < a < b$. We notice that, for $t \in [0, a]$,

$$\begin{aligned}
&\left|\int_0^t \frac{d\hat{G}'(u)}{\hat{S}(u)} - \int_0^t \frac{dG'(u)}{S(u)}\right| \\
&\qquad = \left|\int_0^t \left\{\frac{1}{\hat{S}(u)} - \frac{1}{S(u)}\right\} d\hat{G}'(u) + \int_0^t \frac{1}{S(u)} d\left\{\hat{G}'(u) - G'(u)\right\}\right| \\
&\qquad \le \int_0^t \left|\frac{1}{\hat{S}(u)} - \frac{1}{S(u)}\right| d\hat{G}'(u) + \left|\int_0^t \frac{1}{S(u)} d\left\{\hat{G}'(u) - G'(u)\right\}\right| \\
&\qquad \le \sup_{u\in[0,t]} \left|\frac{1}{\hat{S}(u)} - \frac{1}{S(u)}\right| + \left|\int \frac{\mathbb{I}_{[0,t]}(u)}{S(u)} d\left\{\hat{G}'(u) - G'(u)\right\}\right|
\end{aligned}$$

and therefore, defining $\mathcal{F}_0 = \{f : f(u) = \mathbb{I}_{[0,t]}(u)/S(u) \text{ for some } t \in [0,a] \text{ and all } u\}$, we have that

$$\sup_{t\in[0,a]} \left| \int_0^t \frac{d\hat{G}'(u)}{\hat{S}(u)} - \int_0^t \frac{dG'(u)}{S(u)} \right| \le \frac{\sup_{u\in[0,a]} |\hat{S}(u) - S(u)|}{\hat{S}(a)S(a)} + \sup_{f\in\mathcal{F}_0} \left| \int f(u) d\left\{\hat{G}'(u) - G'(u)\right\} \right|.$$

Both summands converge to zero in probability: the first is a consequence of the uniform consistency of the product-limit estimator $\hat{S}$ (Tsai et al., 1987, and references cited therein) and the second follows from $\mathcal{F}_0$ being a Glivenko-Cantelli class for G'. This last statement is justified by Theorem 19.13 and the first paragraph on page 276 of van der Vaart (2000) along with the fact that $\mathcal{F}_0$ is a Vapnik-Cervonenkis subgraph class, as verified using, for example, Lemma 2.6.16 and part (vi) of Lemma 2.6.18 of van der Vaart & Wellner (2000), with envelope function $F_0(u) = \mathbb{I}_{[0,a]}(u)/S(u)$ satisfying trivially the condition $\int F_0(u)dG'(u) < \infty$. We may write

$$\begin{aligned}\sup_{t\in[0,a]} |\hat{G}(t) - G(t)| &= \sup_{t\in[0,a]} \left| \int_0^t \frac{d\hat{G}'(u)}{\hat{S}(u)}\hat{\beta} - \int_0^t \frac{dG'(u)}{S(u)}\beta \right| \\ &\le |\hat{\beta} - \beta| \int_0^a \frac{d\hat{G}'(u)}{\hat{S}(u)} + \beta \sup_{t\in[0,a]} \left| \int_0^t \frac{d\hat{G}'(u)}{\hat{S}(u)} - \int_0^t \frac{dG'(u)}{S(u)} \right|\end{aligned}$$

and thus conclude, in view of the result above and the consistency of $\hat{\beta}$ (Lemma 4.2 of Wang, 1991), that $\hat{G}$ is uniformly consistent for G on $[0,a]$ for each $0 < a < b$. Now, let $\epsilon > 0$ be given and select $a_\epsilon \in (0,b)$ such that $1 - \epsilon/2 < G(a_\epsilon) < 1$. Because

$$\begin{aligned}\text{pr}\left(\sup_{t\in[a_\epsilon,b]} |\hat{G}(t) - G(t)| \ge \epsilon\right) &\le \text{pr}\left(\max\{1 - \hat{G}(a_\epsilon), 1 - G(a_\epsilon)\} \ge \epsilon\right) \\ &= \text{pr}\left(1 - \hat{G}(a_\epsilon) \ge \epsilon\right) \\ &\le \text{pr}\left(|\hat{G}(a_\epsilon) - G(a_\epsilon)| \ge \epsilon/2\right),\end{aligned}$$

the uniform consistency of $\hat{G}$ on $[0,b]$ follows from the uniform consistency on intervals of the form $[0,a]$ for $0 < a < b$ and the fact that

$$\text{pr}\left(\sup_{t\in[0,b]} |\hat{G}(t) - G(t)| \ge \epsilon\right) \le \text{pr}\left(\sup_{t\in[0,a_\epsilon]} |\hat{G}(t) - G(t)| \ge \epsilon\right) + \text{pr}\left(\sup_{t\in[a_\epsilon,b]} |\hat{G}(t) - G(t)| \ge \epsilon\right).$$

Asymptotic representation of $\{\hat{\Lambda}(t;c) - \Lambda(t;c)\}/n_{\mathbf{pop}}$

For the i^{th} individual in the target population, set the indicator η_i to value one if this individual is a member of the cross-sectional population and to zero otherwise; it follows then that $n_{\mathrm{pop}} = \sum_{i=1}^{n} \eta_i$. Denote by $\omega > 0$ the marginal success probability of these Bernoulli random variables, and observe that $\Phi(\tau) = n\omega\phi(\tau)$. By the Weak Law of Large Numbers, n_{pop}/n converges to ω in probability, and by the Central Limit Theorem and the delta method,

$$n^{1/2}\left(\frac{n}{n_{\mathrm{pop}}} - \frac{1}{\omega}\right)$$

converges to a normal variate. By algebraic manipulation, we may write $r_{n_s}(t)$ as

$$r'_{n_s}(t) + \{\Phi(\tau)/n_{\mathrm{pop}}\}\left\{r''_{n_s}(t) - r'''_{n_s}(t)\right\} + \phi(\tau)\omega\left(\frac{1}{\omega} - \frac{n}{n_{\mathrm{pop}}}\right) r''''_{n_s}(t),$$

where

$$r'_{n_s}(t) = \left\{\frac{\hat{\Phi}(\tau) - \Phi(\tau)}{n_{\mathrm{pop}}}\right\}\left\{\int_{\tau-t}^{\tau-c} \frac{d\hat{G}'(u)}{\hat{S}(u)} - \int_{\tau-t}^{\tau-c} \frac{dG'(u)}{S(u)}\right\},$$

$$r''_{n_s}(t) = \int_{\tau-t}^{\tau-c} \left\{\frac{1}{\hat{S}(u)} - \frac{1}{S(u)}\right\} d\left\{\hat{G}'(u) - G'(u)\right\},$$

$$r'''_{n_s}(t) = \int_{\tau-t}^{\tau-c} \left\{\hat{S}(u) - S(u)\right\}\left\{\frac{1}{\hat{S}(u)} - \frac{1}{S(u)}\right\} \frac{dG'(u)}{S(u)},$$

$$r''''_{n_s}(t) = \int_{\tau-t}^{\tau-c} \frac{dG'(u)}{S(u)} + \int_{\tau-t}^{\tau-c} \left\{\hat{S}(u) - S(u)\right\} \frac{dG'(u)}{S(u)^2} - \int_{\tau-t}^{\tau-c} \frac{1}{S(u)} d\left\{\hat{G}'(u) - G'(u)\right\}.$$

Because

$$\begin{aligned} n_s^{1/2}\left\{\frac{\hat{\Phi}(\tau) - \Phi(\tau)}{n_{\mathrm{pop}}}\right\} &= n_s^{1/2}\left\{\hat{\phi}(\tau) - \phi(\tau)\right\} + \phi\omega\left(\frac{n_s}{n}\right)^{1/2} n^{1/2}\left(\frac{1}{\omega} - \frac{n}{n_{\mathrm{pop}}}\right) \\ &= n_s^{1/2}\left\{\hat{\phi}(\tau) - \phi(\tau)\right\} + o_p(1) \end{aligned}$$

while $n_s^{1/2}\{\hat{\phi}(\tau) - \phi(\tau)\}$ converges to a normal variate, $n_s^{1/2}\{\hat{\Phi}(\tau) - \Phi(\tau)\}/n_{\mathrm{pop}}$ is bounded in probability. This observation, the consistency of $\hat{\beta}$ and the uniform consistency of $\hat{G}$ imply, using the equality

$$r'_{n_s}(t) = \left\{\frac{\hat{\Phi}(\tau) - \Phi(\tau)}{n_{\mathrm{pop}}}\right\}\left\{\frac{\hat{G}(\tau-c) - \hat{G}(\tau-t)}{\hat{\beta}} - \frac{G(\tau-c) - G(\tau-t)}{\beta}\right\},$$

that $\lim_{n_s,n} n_s^{1/2} \sup_{t\in[c,\tau]} |r'_{n_s}(t)| = 0$ in probability. Both integrand and integrator in $r''_{n_s}(t)$ are of bounded variation pathwise, being differences of non-decreasing functions.

Using integration by parts, we may write $r''_{n_s}(t) = r''_{1,n_s}(c) - r''_{1,n_s}(t) - r''_{2,n_s}(t)$, where

$$r''_{1,n_s}(t) = \left\{ \frac{1}{\hat{S}(\tau - t)} - \frac{1}{S(\tau - t)} \right\} \left\{ \hat{G}'(\tau - t) - G'(\tau - t) \right\},$$

$$r''_{2,n_s}(t) = \int_{\tau - t}^{\tau - c} \left\{ \hat{G}'(u) - G'(u) \right\} d \left\{ \frac{1}{\hat{S}(u)} - \frac{1}{S(u)} \right\}.$$

Both $n_s^{1/2} r''_{1,n_s}(c)$ and $n_s^{1/2} \sup_{t \in [c,\tau]} |r''_{1,n_s}(t)|$ converge to zero in probability in view of the uniform consistency of $\hat{S}$ and the fact that $n_s^{1/2} \sup_{u \in [0,\tau]} |\hat{G}'(u) - G'(u)|$ is bounded in probability. Indeed, the empirical process

$$\mathbb{G}_{n_d}(u) = n_d^{1/2} \left\{ \hat{G}'(u) - G'(u) \right\}$$

converges weakly to a tight process $B_2 = B_3 \circ G'$, where B_3 is a Brownian bridge. In particular, B_2 has continuous sample paths almost surely.

Let $\epsilon > 0$ be given. Then, the tightness of B_2 implies that there exists a constant $K = K(\epsilon)$ such that $\mathrm{pr}(\sup_{u \in [0,\tau]} |B_2(u)| \geq K) < \epsilon/2$. Let the operator $\Gamma_\epsilon : D_0 \to D_0$, with D_0 the space of cadlag functions on $[0, \tau]$ vanishing at the endpoints, be defined pointwise as

$$\Gamma_\epsilon B(u) = -K \mathbb{I}_{(-\infty,-K]}(B(u)) + B(u) \mathbb{I}_{[-K,K]}(B(u)) + K \mathbb{I}_{[K,\infty)}(B(u))$$

for each $B \in D_0$ and $u \in [0, \tau]$: this operator has bounded image and is continuous with respect to the Skorohod norm. By the Continuous Mapping Theorem (Theorem 1.11.1 of van der Vaart & Wellner, 1996), $\mathbb{G}_{n_d,\epsilon} = \Gamma_\epsilon \mathbb{G}_{n_d}$ converges weakly to $B_{2,\epsilon} = \Gamma_\epsilon B_2$, a bounded process with continuous sample paths almost surely.

Let $0 < \gamma < S(\tau - c)$ be given and define $\Theta_{n_s}(\gamma)$ to be the subset of the sample space on which $\hat{S}(\tau - c) > S(\tau - c) - \gamma$. The consistency of $\hat{S}$ implies that for some $n(\epsilon) \in \mathbb{N}$, $\mathrm{pr}(\Theta_{n_s}(\gamma)) \geq 1 - \epsilon/2$ holds for each $n_s \geq n(\epsilon)$. Denote

$$\Theta_\epsilon = \cup_{n_s \geq n(\epsilon)} \Theta_{n_s}(\gamma)$$

and note that the complement Θ_ϵ^c of Θ_ϵ satisfies $\mathrm{pr}(\Theta_\epsilon^c) < \epsilon/2$. Define $D_{n_s}(u) = 1/\hat{S}(u) - 1/S(u)$, $D'_{n_s}(u) = D_{n_s}(u) \mathbb{I}_{\Theta_\epsilon}$ and $D''_{n_s}(u) = D_{n_s}(u) \mathbb{I}_{\Theta_\epsilon^c}$. Processes D_{n_s}, D'_{n_s} and D''_{n_s} each converge uniformly to zero in probability on $[0, \tau - c]$. Further, the total variation $\mathcal{V}_c(D'_{n_s})$ of D'_{n_s} over $[0, \tau - c]$ is eventually uniformly bounded: indeed, for $n_s \geq n(\epsilon)$, we have that

$$\mathcal{V}_c(D'_{n_s}) \leq \mathcal{V}_c(\mathbb{I}_{\Theta_\epsilon}/\hat{S}) + \mathcal{V}_c(\mathbb{I}_{\Theta_\epsilon}/S) \leq \left\{ \frac{1}{S(\tau - c) - \gamma} - 1 \right\} + \left\{ \frac{1}{S(\tau - c)} - 1 \right\}$$

using the definition of Θ_ϵ and the elementary fact that the total variation of a monotone function h on an interval $[a_1, a_2]$ is $|h(a_1) - h(a_2)|$. Using Proposition 7.27 and part (iii) of Lemma 7.22 of Kosorok (2008), we find that

$$\sup_{t \in [c,\tau]} \left| \int_{\tau - t}^{\tau - c} \mathbb{G}_{n_d,\epsilon}(u) dD'_{n_s}(u) \right|$$

converges to zero in probability. Further, it is easy to see that, for any $\delta > 0$,

$$\text{pr}\left(\sup_{t\in[c,\tau]}\left|\int_{\tau-t}^{\tau-c}\mathbb{G}_{n_d,\epsilon}(u)dD''_{n_s}(u)\right| > \delta\right) \leq \text{pr}\left(\Theta_\epsilon^c\right) < \epsilon/2\ .$$

Since $D_{n_s} = D'_{n_s} + D''_{n_s}$, we have that $\text{pr}\left(\sup_{t\in[c,\tau]}\left|\int_{\tau-t}^{\tau-c}\mathbb{G}_{n_d,\epsilon}(u)dD_{n_s}(u)\right| > \delta\right)$ is bounded above by

$$\text{pr}\left(\sup_{t\in[c,\tau]}\left|\int_{\tau-t}^{\tau-c}\mathbb{G}_{n_d,\epsilon}(u)dD'_{n_s}(u)\right| > \delta/2\right) + \text{pr}\left(\sup_{t\in[c,\tau]}\left|\int_{\tau-t}^{\tau-c}\mathbb{G}_{n_d,\epsilon}(u)dD''_{n_s}(u)\right| > \delta/2\right)$$

and thus, it is true that $\limsup_{n_s}\text{pr}\left(\sup_{t\in[c,\tau]}\left|\int_{\tau-t}^{\tau-c}\mathbb{G}_{n_d,\epsilon}(u)dD_{n_s}(u)\right| > \delta\right) \leq \epsilon/2$ for any $\delta > 0$. Now, we may finally write

$$\begin{aligned}
\text{pr}\left(n_d^{1/2}\sup_{t\in[c,\tau]}|r''_{2,n_s}(t)| > \delta\right) &= \text{pr}\left(\sup_{t\in[c,\tau]}\left|\int_{\tau-t}^{\tau-c}\mathbb{G}_{n_d}(u)dD_{n_s}(u)\right| > \delta\right)\\
&\leq \text{pr}\left(\sup_{t\in[c,\tau]}\left|\int_{\tau-t}^{\tau-c}\mathbb{G}_{n_d,\epsilon}(u)dD_{n_s}(u)\right| + \sup_{t\in[c,\tau]}\left|\int_{\tau-t}^{\tau-c}\left[\mathbb{G}_{n_d}(u) - \mathbb{G}_{n_d,\epsilon}(u)\right]dD_{n_s}(u)\right| > \delta\right)\\
&\leq \text{pr}\left(\sup_{t\in[c,\tau]}\left|\int_{\tau-t}^{\tau-c}\mathbb{G}_{n_d,\epsilon}(u)dD_{n_s}(u)\right| > \delta/2\right)\\
&\quad + \text{pr}\left(\sup_{t\in[c,\tau]}\left|\int_{\tau-t}^{\tau-c}\left\{\mathbb{G}_{n_d}(u) - \mathbb{G}_{n_d,\epsilon}(u)\right\}dD_{n_s}(u)\right| > \delta/2\right)\\
&\leq \text{pr}\left(\sup_{t\in[c,\tau]}\left|\int_{\tau-t}^{\tau-c}\mathbb{G}_{n_d,\epsilon}(u)dD_{n_s}(u)\right| > \delta/2\right) + \text{pr}\left(\sup_{u\in[0,\tau-c]}|\mathbb{G}_{n_d}(u)| > K\right)
\end{aligned}$$

and therefore, using the results above, we find that

$$\limsup_{n_s}\text{pr}\left(n_d^{1/2}\sup_{t\in[c,\tau]}|r''_{2,n_s}(t)| > \delta\right) \ \leq\ \epsilon/2 + \epsilon/2 \ =\ \epsilon$$

for any $\delta > 0$. Since $\epsilon > 0$ was arbitrary, this proves that $n_d^{1/2}\sup_{t\in[c,\tau]}|r''_{2,n_s}(t)|$ converges to zero in probability and therefore that

$$\lim_{n_s,n} n_s^{1/2}\sup_{t\in[c,\tau]}|r''_{n_s}(t)| = 0$$

in probability.

To show that $\lim_{n_s,n} n_s^{1/2}\sup_{t\in[c,\tau]}|r'''_{n_s}(t)| = 0$ in probability, it suffices to show that

$$\lim_{n_s,n} n_s^{1/2}\sup_{t\in[c,\tau]}\int_{\tau-t}^{\tau-c}\left\{\hat{S}(u) - S(u)\right\}^2\frac{dG'(u)}{S(u)^2} = 0 \tag{1}$$

in probability. Indeed, if $\epsilon > 0$ is given and $0 < \delta < S(\tau - c)$ is selected, we have that

$$\text{pr}\left(n_s^{1/2}\sup_{t\in[c,\tau]}|r'''_{n_s}(t)| > \epsilon\right) \leq \text{pr}\left(n_s^{1/2}\sup_{t\in[c,\tau]}\int_{\tau-t}^{\tau-c}\left\{\hat{S}(u) - S(u)\right\}^2\frac{dG'(u)}{S(u)^2} > \hat{S}(\tau-c)\epsilon\right)$$

$$\leq \mathrm{pr}\left(n_s^{1/2}\sup_{t\in[c,\tau]}\int_{\tau-t}^{\tau-c}\left\{\hat{S}(u)-S(u)\right\}^2\frac{dG'(u)}{S(u)^2}>\hat{S}(\tau-c)\epsilon\ ,\ \hat{S}(\tau-c)>\gamma\right)$$
$$+\mathrm{pr}\left(\hat{S}(\tau-c)\leq\gamma\right)$$
$$\leq \mathrm{pr}\left(n_s^{1/2}\sup_{t\in[c,\tau]}\int_{\tau-t}^{\tau-c}\left\{\hat{S}(u)-S(u)\right\}^2\frac{dG'(u)}{S(u)^2}>\gamma\epsilon\right)+\mathrm{pr}\left(\hat{S}(\tau-c)\leq\gamma\right)$$

and thus prove the desired statement using (1) and the fact that $\mathrm{pr}(\hat{S}(\tau-c)\leq\gamma)\to 0$. That (1) is true follows from the uniform convergence of $\hat{S}$, and the fact that

$$n_s^{1/2}\{\hat{S}(u)-S(u)\}$$

converges weakly to a Gaussian process and is thus uniformly bounded in probability.

Finally, the fact that $\lim_{n_s,n}\Phi(\tau)/n_{\mathrm{pop}}=\phi(\tau)\in(0,1)$ in probability, that $r''''_{n_s}(t)$ is bounded in probability, and that

$$\lim_{n_s,n} n_s^{1/2}\left(\frac{1}{\omega}-\frac{n}{n_{\mathrm{pop}}}\right)=\lim_{n_s,n}\left(\frac{n_s}{n}\right)^{1/2}\left\{n^{1/2}\left(\frac{1}{\omega}-\frac{n}{n_{\mathrm{pop}}}\right)\right\}=0$$

in probability, in conjunction with the results above, suffice to conclude, as required, that

$$\lim_{n_s,n} n_s^{1/2}\sup_{t\in[c,\tau]}|r_{n_s}(t)|=0$$

in probability.

Weak convergence of $\hat{A}(t;c)$

We first establish the marginal weak convergence of each summand in the asymptotic representation to its counterpart in $A(t;c)$. The weak convergence of the first summand is a trivial application of the classical Central Limit Theorem. The second summand is an easy application of the Continuous Mapping Theorem. The process

$$n_d^{1/2}\left\{\hat{S}(u)-S(u)\right\}$$

converges weakly to a mean-zero Gaussian process B_1 (Tsai et al., 1987; page 135 of Wang, 1991; pages 174-177 of Woodroofe, 1985) while $n_s^{1/2}\{\hat{S}(u)-S(u)\}$ converges weakly to $\{\phi(\tau)\}^{-1/2}B_1$, where the scaling factor is the limit in probability of $(n_s/n_d)^{1/2}$. Denote by D_1 the class of cadlag functions on $[0,b]$ vanishing at the endpoints. Define the operator $\mathcal{Q}: D_1\to D_{0c}$ pointwise as $\mathcal{Q}(H)(t)=\int_{\tau-t}^{\tau-c}H(u)dG'(u)/S(u)^2$ for $H\in D_1$ and $t\in[c,\tau]$: this is a linear and bounded operator, with norm at most

$$\int_0^{\tau-c} dG'(u)/S(u)^2<\infty\ .$$

Hence, the Continuous Mapping Theorem (Theorem 1.11.1 of van der Vaart & Wellner, 1996) applies and the weak convergence of the second summand is obtained. For the weak convergence of the third summand, we first observe that

$$n_d^{1/2}\left\{\hat{G}'(u)-G'(u)\right\}$$

converges weakly, as a process, to B_2 by Donsker's Theorem (Theorem 19.3 of van der Vaart, 2000). As before, this implies that $n_s^{1/2}\{\hat{G}'(u) - G'(u)\}$ converges weakly to $\{\phi(\tau)\}^{-1/2}B_2$. In view of Theorem 19.14 and the first paragraph on page 276 of van der Vaart (2000), it then suffices to observe that $\mathcal{F}_1 = \{f : f(u) = \mathbb{I}_{[\tau-t,\tau-c]}(u)/S(u)$ for some $t \in [c,\tau]$ and all $u\}$ is a Donsker class for G', being a Vapnik-Cervonenkis subgraph class with envelope function $F_1(u) = \mathbb{I}_{[0,\tau-c]}(u)/S(u)$ trivially satisfying $\int F_1(u)^2 dG'(u) < \infty$.

As argued in Appendix 3, the limit processes B_0, B_1 and B_2 are orthogonal to each other, and hence, so are the summands in the definition of A. This asymptotic independence suffices, in view of Example 1.4.6 of van der Vaart & Wellner (2000), to conclude that the marginal weak convergence of each summand of the asymptotic representation indeed directly characterizes the weak convergence of the full representation itself.

We now proceed with the calculation of the covariance function for each summand in A. Throughout, fix $s \le t$, both in the interval $[c,\tau]$. We first have that

$$\begin{aligned}
\Sigma_0(s,t) &= E\left[\left\{B_0 \int_{\tau-s}^{\tau-c} \frac{dG'(u)}{S(u)}\right\}\left\{B_0 \int_{\tau-t}^{\tau-c} \frac{dG'(u)}{S(u)}\right\}\right] \\
&= E\left(B_0^2\right) \int_{\tau-s}^{\tau-c} \frac{dG'(u)}{S(u)} \int_{\tau-t}^{\tau-c} \frac{dG'(u)}{S(u)} \\
&= \phi(\tau)\left\{1-\phi(\tau)\right\} \int_{\tau-s}^{\tau-c} \frac{dG'(u)}{S(u)} \int_{\tau-t}^{\tau-c} \frac{dG'(u)}{S(u)} \\
&= \phi(\tau)\left\{1-\phi(\tau)\right\}\left\{G(\tau-c) - G(\tau-s)\right\}\left\{G(\tau-c) - G(\tau-t))\right\}/\beta^2 \,.
\end{aligned}$$

Further, we may write that

$$\begin{aligned}
\Sigma_1(s,t) &= E\left\{\int_{\tau-s}^{\tau-c} B_1(u)\frac{dG'(u)}{S(u)^2} \int_{\tau-t}^{\tau-c} B_1(v)\frac{dG'(v)}{S(v)^2}\right\} \\
&= \int_{\tau-s}^{\tau-c}\int_{\tau-t}^{\tau-c} \Psi(u,v)\frac{dG'(u)}{S(u)^2}\frac{dG'(v)}{S(v)^2} \\
&= \int_{\tau-s}^{\tau-c}\int_{\tau-t}^{\tau-c} \Psi(u,v)\frac{dG(u)}{S(u)}\frac{dG(v)}{S(v)}/\beta^2 \\
&= \int_{\tau-s}^{\tau-c}\int_{\tau-t}^{\tau-c}\int_0^{u\wedge v} \frac{dF(x)}{R(x)S(x)} dG(u)dG(v)/\beta^2 \,.
\end{aligned}$$

With obvious abuse of notation, the trivariate integration operator above may be simplified by using Fubini's Theorem:

$$\begin{aligned}
\int_{v=\tau-s}^{\tau-c}\int_{u=\tau-t}^{\tau-c}\int_{x=0}^{u\wedge v} &= \int_{v=\tau-s}^{\tau-c}\int_{u=\tau-t}^{v}\int_{x=0}^{u} + \int_{v=\tau-s}^{\tau-c}\int_{u=v}^{\tau-c}\int_{x=0}^{v} \\
&= \int_{v=\tau-s}^{\tau-c}\int_{x=0}^{v}\int_{u=(\tau-t)\vee x}^{v} + \int_{v=\tau-s}^{\tau-c}\int_{x=0}^{v}\int_{u=v}^{\tau-c} \\
&= \int_{x=0}^{\tau-c}\int_{v=(\tau-s)\vee x}^{\tau-c}\int_{u=(\tau-t)\vee x}^{v} + \int_{x=0}^{\tau-c}\int_{v=(\tau-s)\vee x}^{\tau-c}\int_{u=v}^{\tau-c} \\
&= \int_{x=0}^{\tau-c}\int_{v=(\tau-s)\vee x}^{\tau-c}\int_{u=(\tau-t)\vee x}^{\tau-c} \,.
\end{aligned}$$

Therefore, we obtain $\Sigma_1(s,t)$ to be

$$\int_{x=0}^{\tau-c}\int_{v=(\tau-s)\vee x}^{\tau-c}\int_{u=(\tau-t)\vee x}^{\tau-c} dG(u)dG(v)\frac{dF(x)}{R(x)S(x)}/\beta^2$$
$$= \int_0^{\tau-c}\frac{[G(\tau-c)-G\{(\tau-s)\vee x\}]\,[G(\tau-c)-G\{(\tau-t)\vee x\}]}{R(x)S(x)}dF(x)/\beta^2 \ .$$

Finally, because we may write that

$$\int_{\tau-t}^{\tau-c}\frac{1}{S(u)}d\left[n_d^{1/2}\left\{\hat{G}'(u)-G'(u)\right\}\right] = n_d^{1/2}\left\{\frac{1}{n_d}\sum_{i=1}^{n_d}\frac{\mathbb{I}_{[\tau-t,\tau-c]}(T_i)}{S(T_i)}-\int_{\tau-t}^{\tau-c}\frac{dG'(u)}{S(u)}\right\} \ ,$$

the classical Central Limit Theorem (Example 2.1 of van der Vaart, 2000) may be used to conclude that

$$\begin{aligned}
\Sigma_2(s,t) &= E\left\{\frac{\mathbb{I}_{[\tau-s,\tau-c]}(T)}{S(T)}\frac{\mathbb{I}_{[\tau-t,\tau-c]}(T)}{S(T)}\right\} - E\left\{\frac{\mathbb{I}_{[\tau-s,\tau-c]}(T)}{S(T)}\right\}E\left\{\frac{\mathbb{I}_{[\tau-t,\tau-c]}(T)}{S(T)}\right\}\\
&= \int_{\tau-s}^{\tau-c}\frac{dG'(u)}{S(u)^2} - \int_{\tau-s}^{\tau-c}\frac{dG'(u)}{S(u)}\int_{\tau-t}^{\tau-c}\frac{dG'(u)}{S(u)}\\
&= \int_{\tau-s}^{\tau-c}\frac{dG(u)}{S(u)}/\beta - \{G(\tau-c)-G(\tau-s)\}\{G(\tau-c)-G(\tau-t)\}/\beta^2 \ .
\end{aligned}$$

Supplementary material

Additional simulation results

Table 4 provides, for each simulation scenario considered in Section 4, the empirical standard error of $\hat{\Lambda}(t;10)$ at $t=20,60,100$ and for effective sample size 250, 500 and 1,000. These estimates were obtained by simulating 5,000 datasets for each combination of parameters considered. Table 4 is obtained from the same simulation study reported in Tables 1 and 2.

Simulation studies were performed to investigate the sampling distribution of $\hat{\Lambda}(t;10)$ and its discrepancy with respect to the family of normal distributions. For scenarios I, II, III and IV, and for effective sample size 250, 500 and 1,000, more than 1,000 datasets were generated, and $\hat{\Lambda}(t;10)$ was computed at $t=30,70$. Figs. 4–7 are quantile-quantile plots assessing the normality of the sampling distribution of interest.

Additional simulation studies were also conducted to illustrate the estimation of age-specific incidence rates using the estimator presented in Subsection 3.2. In the following, we denote by Z^0, $W^0=Z^0+X^0$ and D^0, respectively, age at disease onset, age at death, and lifetime disease status, defined as the binary indicator of ever developing the disease of interest. These simulation studies were carried out by assuming the following working structure, which rendered the computation of theoretical quantities mathematically tractable:

- births in the population arise from a stationary Poisson process with rate $\lambda_0>0$;
- lifetime disease status D^0 is independent of date of birth and is a Bernoulli random variable with success probability $\pi\in(0,1)$;

- the law of age at disease onset Z^0 given $D^0 = 1$ has support $\mathcal{Z}$;
- for some partition $\mathcal{P}_{\mathcal{Z}} = (\mathscr{Z}_1, ..., \mathscr{Z}_k)$ of $\mathcal{Z}$, the law of X^0 given $Z^0 = z$ and $D^0 = 1$ is identical to that of X^0 given $Z^0 \in \mathscr{Z}_j$ and $D^0 = 1$, provided $\{z\} \subset \mathscr{Z}_j$;

Denoting by p_j the probability pr $\left(Z^0 \in \mathscr{Z}_j \mid D^0 = 1\right)$, it is not difficult to show that, under the above structure, the average number of individuals of age at least m in the population is, at any given time,

$$n_{pop} = \lambda_0 \Bigg[\pi \left\{ \mathrm{E}\left(Z^0 - m \mid D^0 = 1\right) + \sum_{j=1}^{k} \mathrm{E}\left(X^0 \mid Z^0 \in \mathscr{Z}_j, D^0 = 1\right) p_j \right\} + (1-\pi)\mathrm{E}\left(W^0 - m \mid W^0 > m, D^0 = 0\right) \mathrm{pr}(W^0 > m \mid D^0 = 0) \Bigg] ,$$

provided m is at most the infimum of $\mathcal{Z}$. Furthermore, it is easy to see that the age-specific incidence rate for age-group $\mathscr{Z}_j$ is constant through time with value

$$\lambda_j = \lambda_0 \pi p_j$$

and that the mean prevalence of age-specific disease for age-group $\mathscr{Z}_j$ is itself constant at

$$P_j = \lambda_0 \pi p_j \mathrm{E}\left(X^0 \mid Z^0 \in \mathscr{Z}_j, D^0 = 1\right) .$$

These theoretical quantities were used to validate results from our simulation studies.

Simulation outputs are presented below for the above working structure with particular distributions and parameter values according to the follow specifications:

1. fix $\lambda_0 = 300{,}000$ births per year;
2. set $\pi = 0.4$;
3. let Z^0 given $D^0 = 1$ have the law of $60 + 40B_z$, where B_z is a beta variate with parameters $\alpha = \beta = 2$;
4. define $\mathscr{Z}_j = [60 + 10(j-1), 60 + 10j)$ for $j = 1, ..., 4$, which forms a partition of $\mathcal{Z} = [60, 100)$, and let X^0 given $Z^0 \in \mathscr{Z}_j$ and $D^0 = 1$ have the law of $20B_y$, where B_y is a beta variate with parameters as listed below:

j	α	β
1	12	5
2	10	10
3	8	20
4	4	30

5. let W^0 given $D^0 = 0$ have the law of $60 + 60B_w$, where B_w is a beta variate with parameters $\alpha = 3$ and $\beta = 10$.

Fig. 8 consists of an overlay of the age-specific cumulative incidence estimated from each of 50 random datasets generated according to the above guidelines. The size of the cross-sectional sample was set at 20,000; this corresponds, on average, to approximately 905, 1,410, 806 and 151 sampled prevalent cases from age-subgroups 60-70, 70-80, 80-90 and 90-100, respectively. The age-specific curves were trimmed at $t = 84, 87, 92, 97,$

Table 4. *Estimated standard error of* $\hat{\Lambda}(t;10)$ *at* $t = 20, 60, 100$ *and for effective sample size* $E(n_d) = 250, 500, 1000$ *under scenarios I, II, III and IV using 5,000 random datasets*

t	20	20	20	60	60	60	100	100	100
$E(n_d)$	250	500	1,000	250	500	1,000	250	500	1,000
Scenario									
I	4,958	3,457	2,482	6,790	4,823	3,405	7,895	5,565	4,008
II	3,504	2,434	1,717	5,361	3,783	2,665	7,000	4,927	3,486
III	12,798	8,454	5,966	29,124	20,153	13,343	40,350	28,871	18,611
IV	8,061	5,481	3,847	18,286	12,637	8,927	28,196	20,012	14,196

respectively. Fig. 8 provides an illustration of the appropriateness of the extension of our methodology to the estimation of age-specific rates.

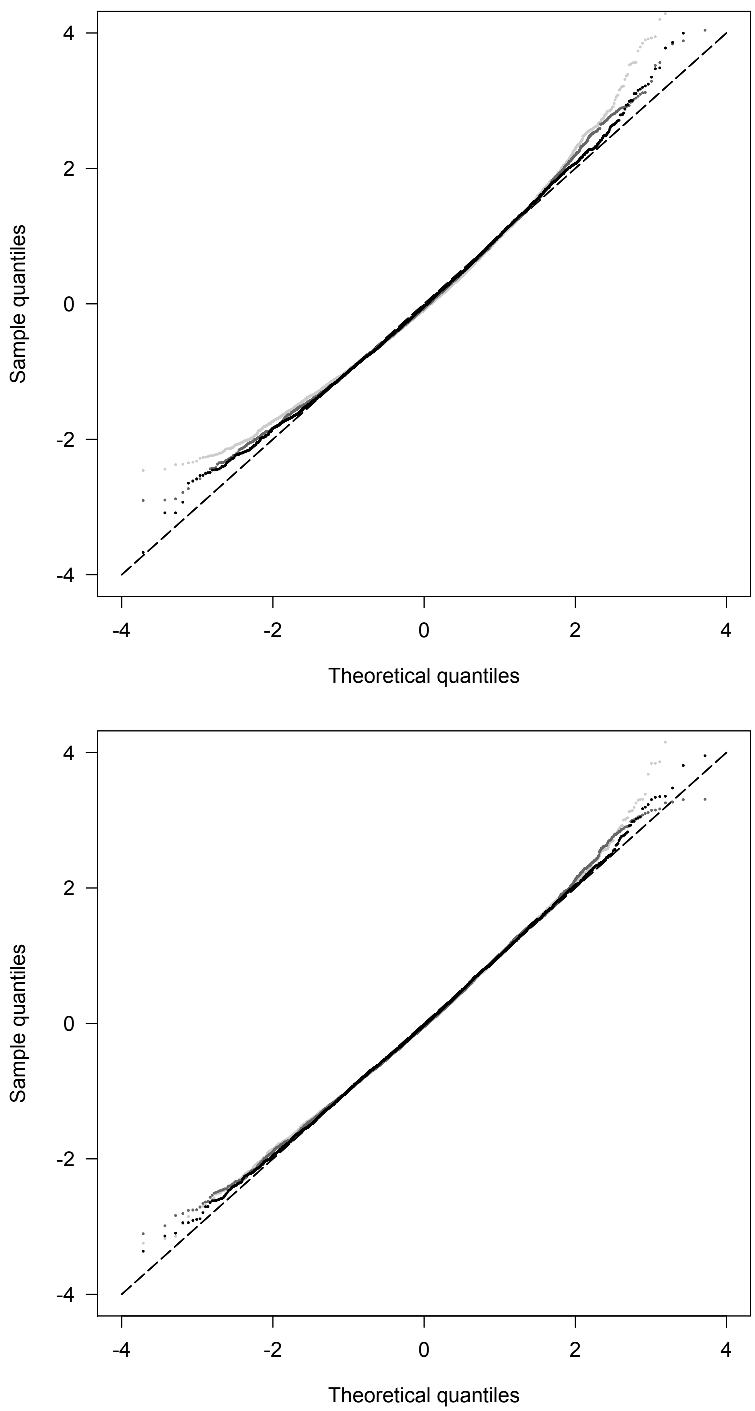

Fig. 4. Quantile-quantile plot assessing the normality of the sampling distribution of $\hat{\Lambda}(t; 10)$ at $t = 30$ (top) and 70 (bottom) and for effective sample size 250 (light grey dots), 500 (grey dots) and 1000 (black dots) under scenario I.

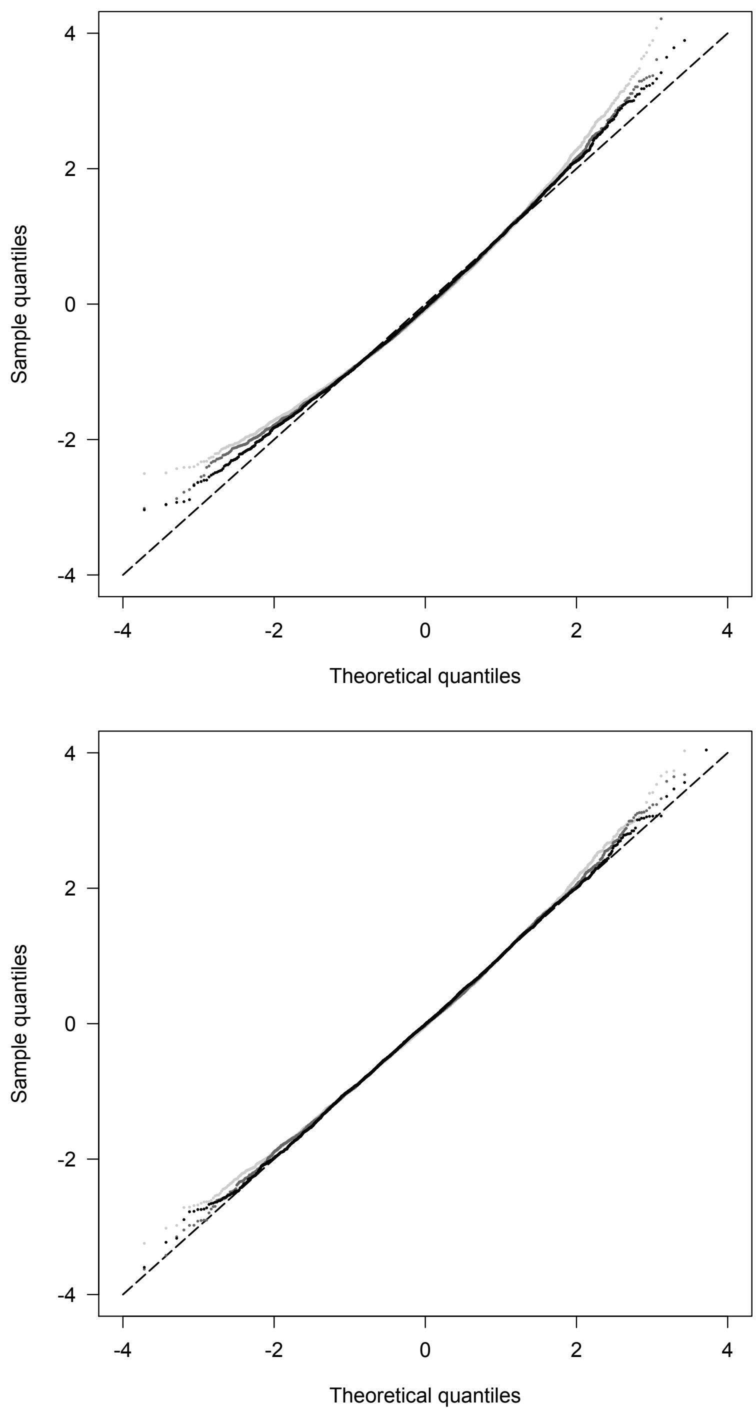

Fig. 5. Quantile-quantile plot assessing the normality of the sampling distribution of $\hat{\Lambda}(t;10)$ at $t = 30$ (top) and 70 (bottom) and for effective sample size 250 (light grey dots), 500 (grey dots) and 1000 (black dots) under scenario II.

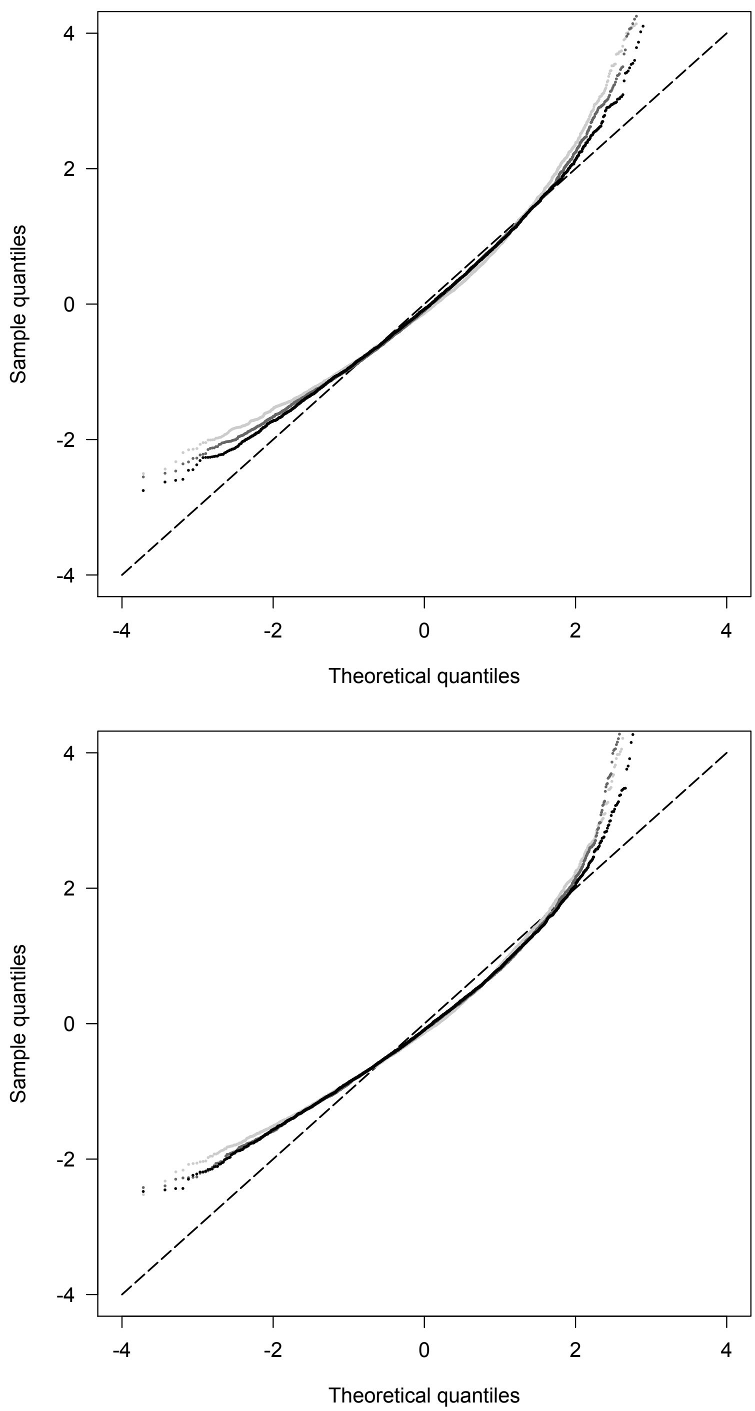

Fig. 6. Quantile-quantile plot assessing the normality of the sampling distribution of $\hat{\Lambda}(t;10)$ at $t = 30$ (top) and 70 (bottom) and for effective sample size 250 (light grey dots), 500 (grey dots) and 1000 (black dots) under scenario III.

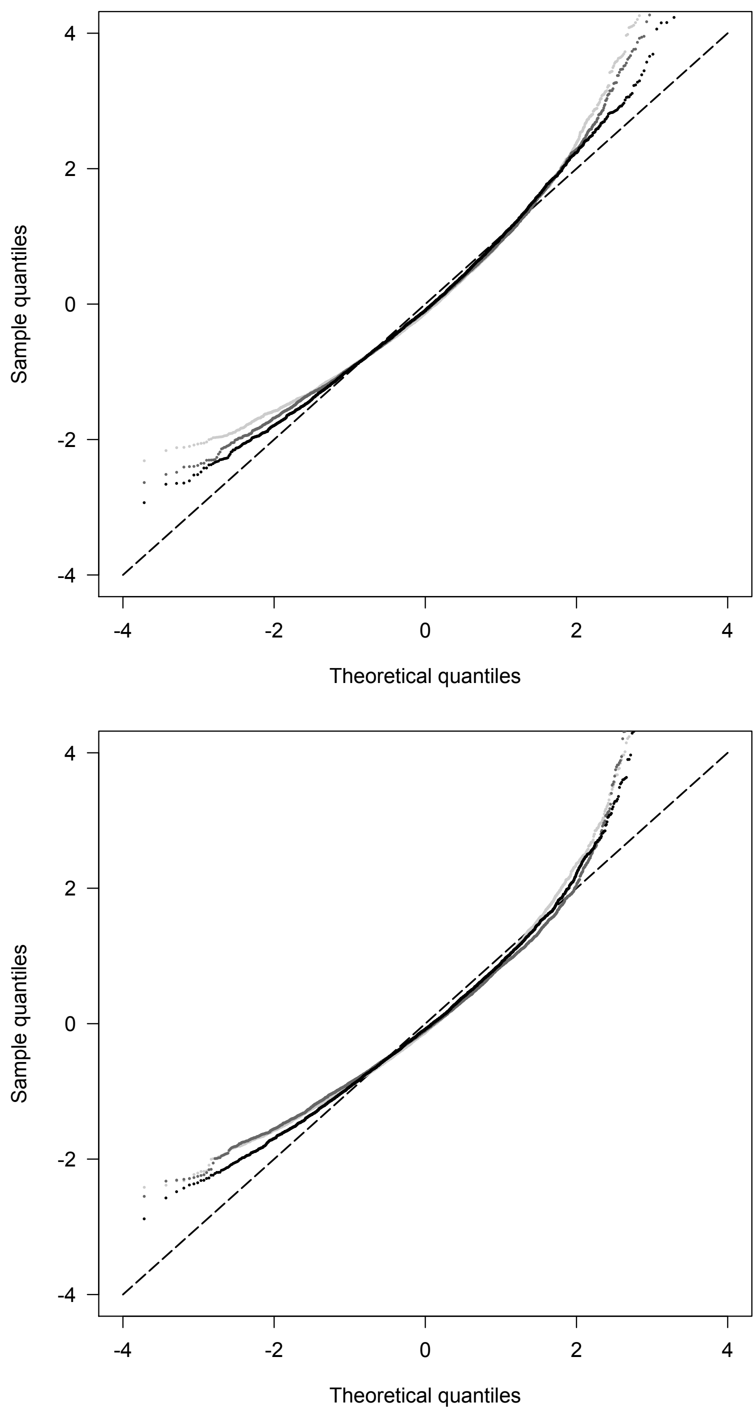

Fig. 7. Quantile-quantile plot assessing the normality of the sampling distribution of $\hat{\Lambda}(t;10)$ at $t = 30$ (top) and 70 (bottom) and for effective sample size 250 (light grey dots), 500 (grey dots) and 1000 (black dots) under scenario IV.

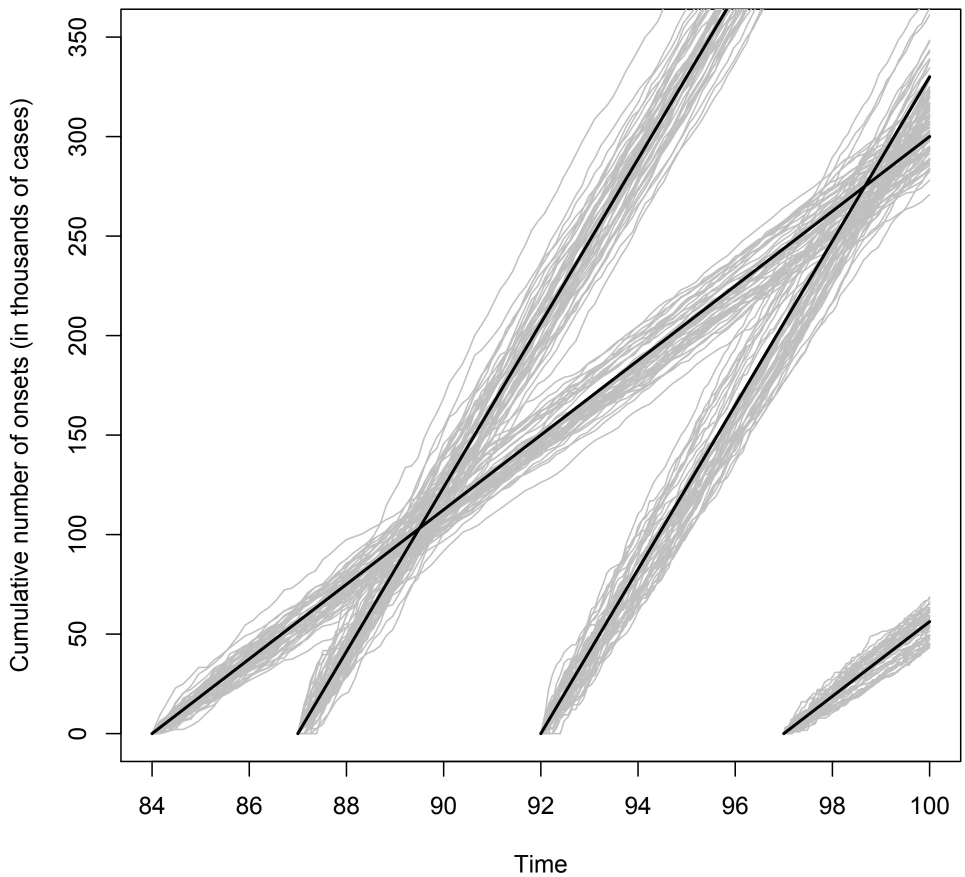

Fig. 8. Overlay of 50 plots of the estimated age-specific cumulative incidence. From left to right, curves correspond to age-groups 60-70, 70-80, 80-90 and 90-100, respectively. Grey lines represent estimated curves and black lines indicate true theoretical curves.