where $h : [0, \infty) \to R$ is a bounded function. Note that $\sigma_F(\cdot)[h]$ is the Fisher information (Murphy 1995, p. 189) with h being the index of infinite-dimensional parameters. Appendix B1 shows that $\sqrt{n}(\hat{F}(t) - F(t))$ converges weakly to a Gaussian process $G_F(t)$ with $E[G_F(t)] = 0$ and

$$E[G_F(s)G_F(t)] = \int w_s(x)\sigma_F^{-1}(w_t)(x)dF(x),$$

where $w_s(x) \equiv \mathbf{I}(x \le s)$ and $\sigma_F^{-1}(w_t)$ solves $\sigma_F(x)[h] = w_t(x)$ for h.

Consider the empirical estimator of $\sigma_F(x)[h]$ as

$$\hat{\sigma}_F(x)[h] = \frac{1}{n}\sum_{i=1}^{n} \mathbf{I}(U_i \le x \le V_i)\left\{\frac{1}{\hat{F}_i}h(x) - \frac{1}{\hat{F}_i^2}\sum_{k=1}^{n} J_{ik}h_k\hat{f}_k\right\},$$

where $h_k = h(T_k)$. Then, the plug-in covariance estimator is

$$\hat{E}[G_F(s)G_F(t)] = \int w_s(x)\hat{\sigma}_F^{-1}(w_t)(x)d\hat{F}(x) = \sum_{j=1}^{n} w_s(T_j)\hat{\sigma}_F^{-1}(w_t)(T_j)\hat{f}_j.$$

After some matrix calculations given in Appendix B2, one can verify

$$\sum_{j=1}^{n} w_s(T_j)\hat{\sigma}_F^{-1}(w_t)(T_j)\hat{f}_j = \mathbf{W}_s^{\mathrm{T}}\left\{\frac{i_n(\hat{\mathbf{f}})}{n}\right\}^{-1}\mathbf{W}_t,$$

where $\mathbf{W}_t = (\mathbf{I}(T_{(1)} \le t) - \mathbf{I}(T_{(n)} \le t), \ldots, \mathbf{I}(T_{(n-1)} \le t) - \mathbf{I}(T_{(n)} \le t))^{\mathrm{T}}$ and $i_n(\mathbf{f})$ is given in Eq. (2). Therefore, we obtain a plug-in covariance estimator

$$\hat{C}ov\{\hat{F}(s), \hat{F}(t)\} = \mathbf{W}_s^{\mathrm{T}}\left[D\left\{\mathrm{diag}\left(\frac{1}{\hat{\mathbf{f}}^2}\right) - J^{\mathrm{T}}\mathrm{diag}\left(\frac{1}{\hat{\mathbf{F}}^2}\right)J\right\}D^{\mathrm{T}}\right]^{-1}\mathbf{W}_t, \quad (3)$$

and a variance estimator

$$\hat{V}_{\mathrm{Info}}\{\hat{F}(t)\} = \mathbf{W}_t^{\mathrm{T}}\left[D\left\{\mathrm{diag}\left(\frac{1}{\hat{\mathbf{f}}^2}\right) - J^{\mathrm{T}}\mathrm{diag}\left(\frac{1}{\hat{\mathbf{F}}^2}\right)J\right\}D^{\mathrm{T}}\right]^{-1}\mathbf{W}_t. \quad (4)$$

Remark: Murphy (1995), Zeng and Lin (2006), Chen (2010), and Emura and Wang (2012) use similar techniques to derive variance estimators. However, none of them results in an explicit form like Eqs. (3) and (4).

4 Inference based on the asymptotic covariance estimator

This section examines various inference procedures based on the proposed covariance estimator.

4.1 Pointwise confidence interval

Applying the variance estimator $\hat{V}_{\text{Info}}\{\hat{F}(t)\}$ in Eq. (4) and the asymptotic normality, we propose a pointwise confidence interval. Log-transformation and arcsine-square root transformation are known to improve the coverage performance over the linear confidence interval (Klein and Moeschberger 2003, pp. 104–108). Here, we apply the log-transformed interval based on $\log \hat{F}(t) - \log F(t) \sim N(0,\ \hat{V}_{\text{Info}}\{\hat{F}(t)\}/\hat{F}(t)^2)$. Hence, the $(1-\alpha)100\%$ confidence interval for $F(t)$ is

$$\left(\hat{F}(t)\exp[-z_{\alpha/2}\hat{V}_{\text{Info}}^{1/2}\{\hat{F}(t)\}/\hat{F}(t)],\quad \hat{F}(t)\exp[z_{\alpha/2}\hat{V}_{\text{Info}}^{1/2}\{\hat{F}(t)\}/\hat{F}(t)]\right),$$

where $z_{\alpha/2}$ is the $(1-\alpha/2)100\%$ point of the standard normal distribution.

4.2 Goodness-of-fit test

We consider a goodness-of-fit test for

$$H_0:\ F = F_0 \text{ vs. } H_1:\ F \neq F_0,$$

where F_0 is a known continuous distribution function. Applying the continuous mapping theorem to the results of Sect. 3.2, we have

$$\sqrt{n}\sup_t|\hat{F}(t) - F(t)| \xrightarrow{d} \sup_t|G_F(t)|.$$

The asymptotic distribution can be easily simulated after estimating the covariance structure of $G_{F_0}(t)$ with Eq. (3). Ideally, the asymptotic distribution is approximated by that of $\max_j |G_{F_0}(t_j)|$ for fixed fine grids $t_j : j = 1, \ldots, N$ with large N. Here, we suggest a practically convenient choice of $t_j = T_{(j)}$, $j = 1, \ldots, n-1$, which leads to a simple algorithm and achieves good finite sample performance. The algorithm is stated as follows:

Kolmogorov–Smirnov test for $H_0:\ F = F_0$ vs. $H_1:\ F \neq F_0$;

Step 1: Calculate $K = \sup_t|\hat{F}(t) - F_0(t)|$ and $i_n(\hat{\mathbf{f}})$.

Step 2: Generate $\mathbf{G}^{(b)} = (G_1^{(b)}, \ldots, G_{n-1}^{(b)}) \sim N(\mathbf{0}_{n-1}, Hi_n(\hat{\mathbf{f}})^{-1}H^{\mathrm{T}})$ for $b = 1, \ldots, B$, and compute $K^{(b)} = \max_{i=1,\ldots,n-1}|G_i^{(b)}|$, where $H = (\mathbf{W}_{T_{(1)}}, \ldots, \mathbf{W}_{T_{(n-1)}})^{\mathrm{T}}$.

Step 3: Reject $H_0:\ F = F_0$ with level α if $\sum_{b=1}^{B}\mathbf{I}(K^{(b)} > K)/B < \alpha$.

Similarly, we can test $H_0:\ F = F_0$ using the Cramér–von Mises statistic

$$C = n\int_0^\infty \{\hat{F}(t) - F_0(t)\}^2 dF_n(t) = \sum_{j=1}^{n}\{\hat{F}(T_j) - F_0(T_j)\}^2,$$

where $F_n(t) = \sum_{j=1}^{n}\mathbf{I}(T_j \le t)/n$ is the empirical distribution function.

Cramér–von Mises test for $H_0: \ F = F_0$ vs. $H_1: \ F \neq F_0$;

Step 1: Calculate $C = \sum_{j=1}^{n} \{ \hat{F}(T_j) - F_0(T_j) \}^2$ and $i_n(\hat{\mathbf{f}})$.
Step 2: Generate $\mathbf{G}^{(b)} = (G_1^{(b)}, \ \ldots, \ G_{n-1}^{(b)}) \sim N(\mathbf{0}_{n-1}, \ H i_n(\hat{\mathbf{f}})^{-1} H^{\mathrm{T}})$ for $b = 1, \ \ldots, \ B$, and then compute $C^{(b)} = (\mathbf{G}^{(b)})^{\mathrm{T}} \mathbf{G}^{(b)}$.
Step 3: Reject $H_0: \ F = F_0$ with level α if $\sum_{b=1}^{B} \mathbf{I}(\ C^{(b)} > C \ )/B < \alpha$.

4.3 Confidence band

The confidence band covers the true function $F(t)$ at all t for a specified confidence level $(1 - \alpha)$. We follow the construction of two most well-known confidence bands for the survival function under right-censoring, namely, the equal precision (EP) band and Hall–Wellner (HW) band (Nair 1984; Klein and Moeschberger 2003, Sect. 4.4).

Let $\psi(u)$ be a nonnegative continuous function. Applying the continuous mapping theorem to the results of Sect. 3.2, we have

$$\sqrt{n} \sup_t | \ \psi\{ \ F(t) \ \}\{ \ \hat{F}(t) - F(t) \ \}| \xrightarrow{d} \sup_t |\psi\{ \ F(t) \ \} G_F(t)|.$$

Then, the confidence bands are obtained by solving

$$1 - \alpha = \Pr\left\{ \sup_t | \ \psi\{ \ F(t) \ \}\{ \ \hat{F}(t) - F(t) \ \}| \ \leq c_{1-\alpha}(\psi) \right\},$$

where $c_{1-\alpha}(\psi)$ is the $(1 - \alpha)100\%$ point of $\sup_t |\psi\{ \ F(t) \ \} G_F(t)/\sqrt{n}|$.

The EP band corresponds to $\psi(u) = \{ \ u(\ 1 - u \ ) \ \}^{-1/2}$. In practice, it is desirable to make $\psi(u)$ bounded. Following Nair (1984), we alternatively use $\psi_{EP}(u) = \{ \ (\ u \vee p_1 \ )(\ 1 - u \wedge p_2 \ ) \ \}^{-1/2}$, $0 < p_1 < p_2 < 1$, to yield the EP band

$$\hat{F}(t) \pm c_{1-\alpha}(\psi_{EP}) \sqrt{\{ \ \hat{F}(t) \vee p_1 \ \}\{ \ 1 - \hat{F}(t) \wedge p_2 \ \}}.$$

We set p_1 =0.1 or 0.2 and p_2 =0.8 or 0.9 as suggested by Nair (1984).

The HW band corresponds to $\psi_{HW}(u) \equiv 1$, which is the version of Kolmogorov–Smirnov band for uncensored data. The HW band is

$$\hat{F}(t) \pm c_{1-\alpha}(\psi_{HW}),$$

where $c_{1-\alpha}(\psi_{HW})$ is obtained as the $(1 - \alpha)100\%$ point for $\{ \ K^{(b)}; \ b = 1, \ \ldots, \ B \ \}$ in Step 2 of the Kolmogorov–Smirnov test.

The bootstrap is useful to validate the coverage performance of the confidence bands above. First, the bootstrap NPMLEs, denoted as $\{ \ \hat{F}_b^*, \ b = 1, \ \ldots, \ B \ \}$, are computed (see Step 1 of Appendix A). Then, approximately $(1 - \alpha)100\%$ of the bootstrap NPMLEs should fall inside the band. This validation scheme will be illustrated with real data analysis.

5 Simulations

Extensive simulations have been conducted to investigate the performances of the proposed methods and to compare them with the bootstrap and jackknife methods.

We adopt the same design used in Moreira and Uña-Álvarez (2010). They consider models $U^* \sim \text{Unif}(0,\ a)$, $T^* \sim \text{Unif}(0,\ 1)$, and $V^* \sim \text{Unif}(b,\ 1)$, where $(a,\ b) = (0.25,\ 0.75)$, $(0.5,\ 0.5)$ or $(0.67,\ 0.33)$. The corresponding sample inclusion probabilities are $\Pr(U^* \leq T^* \leq V^*) = (1 - a + b)/2$ =0.75, 0.5 and 0.33, respectively. They also consider a model $U^* \sim \text{Unif}(-5,\ 15)$, $T^* \sim \text{Unif}(0,\ 15)$, and $V^* = U^* + c$, where c =5. This model is important since it yields a situation similar to the childhood cancer example.

Based on simulated samples, we compute the relevant quantities (NPMLE, confidence interval, goodness-of-fit statistic, and confidence band) for 500 repetitions. We choose B = 1000 for the number of resamplings.

5.1 Performance of the covariance estimator

For r (= 1, ..., 500)-th repetition, we compute the NPMLE $\hat{F}(s)_{(r)}$, $\hat{F}(t)_{(r)}$ and the covariance estimator $\hat{C}ov\{\ \hat{F}(s),\ \hat{F}(t)\ \}_{(r)}$ in Eq. (3). We compare the average of the estimated covariance

$$\frac{1}{500}\sum_{r=1}^{500} \hat{C}ov\left\{\ \hat{F}(s),\ \hat{F}(t)\right\}_{(r)}$$

with the sample covariance

$$\frac{1}{500}\sum_{r=1}^{500}\left\{\ \hat{F}(s)_{(r)} - \bar{\hat{F}}(s)\right\}\left\{\ \hat{F}(t)_{(r)} - \bar{\hat{F}}(t)\right\}.$$

where $\bar{\hat{F}}(s) = \sum_{r=1}^{500} \hat{F}(s)_{(r)}/500$. As shown in Table 1, the differences between the estimated covariance and the sample covariance are very small for all configurations. The sample covariance between $\hat{F}(s)_{(r)}$ and $\hat{F}(t)_{(r)}$ increases as the distance $|t - s|$ decreases, which is a similar behavior to that of the empirical distribution function from un-truncated data.

5.2 Comparison with the bootstrap and jackknife methods

We compare the performance of the proposed variance estimator ($\hat{V}_{\text{Info}}\{\hat{F}(t)\}$), the bootstrap estimator ($\hat{V}_{\text{Boot}}\{\hat{F}(t)\}$) and the jackknife estimator ($\hat{V}_{\text{Jack}}\{\hat{F}(t)\}$) for fixed t. We compute the average of the estimated standard deviation (SD)

$$\frac{1}{500}\sum_{r=1}^{500}\sqrt{\hat{V}\{\hat{F}(t)\}_{(r)}}$$

Table 1 Simulation results of the proposed covariance estimator based on 500 replications

		$n = 100$		$n = 250$	
	$(s,\ t)$	Sample covariance	Estimated covariance	Sample covariance	Estimated covariance
a= 0.25, b = 0.75	(0.2, 0.5)	0.00565	0.00595	0.00195	0.00180
	(0.4, 0.5)	0.00655	0.00668	0.00234	0.00215
	(0.2, 0.8)	0.00279	0.00366	0.00112	0.00110
a= 0.5, b = 0.5	(0.2, 0.5)	0.00641	0.00704	0.00266	0.00258
	(0.4, 0.5)	0.00833	0.00877	0.00335	0.00336
	(0.2, 0.8)	0.00378	0.00450	0.00184	0.00169
a= 0.067, b = 0.33	(0.2, 0.5)	0.00366	0.00374	0.00157	0.00141
	(0.4, 0.5)	0.00677	0.00645	0.00283	0.00245
	(0.2, 0.8)	0.00356	0.00350	0.00159	0.00130
c = 5	(3.0, 7.5)	0.01166	0.01291	0.00525	0.00511
	(6.0, 7.5)	0.02027	0.02068	0.00860	0.00845
	(3.0, 12.0)	0.00558	0.00654	0.00269	0.00260

Data are generated from $U^* \sim$ Unif(0, a), $T^* \sim$ Unif(0, 1), and $V^* \sim$ Unif(b, 1) in the first three cases, and from $U^* \sim$ Unif(-5, 15), $T^* \sim$ Unif(0, 15), and $V^* = U^* + c$ in the last case

Sample covariance = $\frac{1}{500}\sum_{r=1}^{500}\{\hat{F}(s)_{(r)} - \bar{\hat{F}}(s)_{(r)}\}\{\hat{F}(t)_{(r)} - \bar{\hat{F}}(t)_{(r)}\}$

Estimated covariance = $\frac{1}{500}\sum_{r=1}^{500}\hat{C}ov\{\hat{F}(s),\ \hat{F}(t)\}_{(r)}$

where $\hat{V}\{\hat{F}(t)\}_{(r)}$ is a variance estimator for the rth repetition, and compare it with $\mathrm{SD}\{\hat{F}(t)\}$, the sample standard deviation (SD) for $\hat{F}(t)_{(r)}$, $r = 1, \ldots, 500$. The performance of the three methods are measured with the mean squared error

$$\mathrm{MSE} = \frac{1}{500}\sum_{r=1}^{500}\left(\sqrt{\hat{V}\{\hat{F}(t)\}_{(r)}} - \mathrm{SD}\{\hat{F}(t)\}\right)^2.$$

We also compare the performance of the three methods in terms of the coverage performance of the 95% confidence interval.

Tables 2 and 3 show the results under the models $U^* \sim$ Unif(0, a), $T^* \sim$ Unif(0, 1), and $V^* \sim$ Unif(b, 1), where $(a,\ b) = (0.25,\ 0.75)$ and $(0.5,\ 0.5)$, respectively. All the three variance estimators correctly capture the estimates of $\mathrm{SD}\{\hat{F}(t)\}$. Among the three estimators, the jackknife has the smallest bias. In terms of MSE, the bootstrap is the best for small samples, while the proposed method tends to be the best for large samples. For instance, the bootstrap is the best for $n = 100$, while the proposed method is the best for n =200, 250 and 300 (Table 2). The jackknife has the largest MSE in most configurations.

All the three methods generally produce the nominal 95% coverage performance at $t = 0.5$ ($F(t) = 0.5$). However, at the tail $t = 0.2$ ($F(t) = 0.2$), the bootstrap method often results in serious under-coverage. The magnitude of the under-coverage of the bootstrap is similar to that reported in the simulation results of Moreira and Uña-Álvarez (2010). Both the proposed and the jackknife methods alleviate the under-

Table 2 Simulation results under $U^* \sim$ Unif(0, a), $T^* \sim$ Unif(0, 1), and $V^* \sim$ Unif(b, 1) with a= 0.25 and b = 0.75 based on 500 replications

		n = 100	n = 150	n = 200	n = 250	n = 300
$F(t) = 0.5$						
SD		0.083	0.064	0.053	0.050	0.046
ESD	Proposed	0.070	0.057	0.050	0.045	0.042
	Bootstrap	0.070	0.059	0.051	0.046	0.043
	Jackknife	0.075	0.061	0.053	0.047	0.044
MSE	Proposed	0.00219	0.00086	0.00033	0.00028	0.00026
	Bootstrap	0.00104	0.00070	0.00048	0.00038	0.00035
	Jackknife	0.00296	0.00185	0.00094	0.00075	0.00073
95%Cov	Proposed	0.930	0.942	0.950	0.946	0.938
	Bootstrap	0.920	0.938	0.950	0.942	0.948
	Jackknife	0.930	0.950	0.948	0.946	0.940
$F(t) = 0.2$						
SD		0.090	0.065	0.057	0.052	0.048
ESD	Proposed	0.069	0.056	0.049	0.045	0.041
	Bootstrap	0.069	0.058	0.051	0.046	0.042
	Jackknife	0.074	0.061	0.053	0.047	0.043
MSE	Proposed	0.00394	0.00091	0.00073	0.00052	0.00044
	Bootstrap	0.00213	0.00113	0.00094	0.00067	0.00055
	Jackknife	0.00522	0.00248	0.00189	0.00115	0.00103
95%Cov	Proposed	0.938	0.942	0.946	0.932	0.942
	Bootstrap	0.898	0.910	0.928	0.908	0.924
	Jackknife	0.940	0.948	0.952	0.938	0.950

ESD = $\frac{1}{500}\sum_{r=1}^{500}\sqrt{\hat{V}\{\hat{F}(t)\}_{(r)}}$

MSE = $\frac{1}{500}\sum_{r=1}^{500}\left(\sqrt{\hat{V}\{\hat{F}(t)\}_{(r)}} - \mathrm{SD}\{\hat{F}(t)\}\right)^2$

95%Cov = Empirical coverage probability of the 95% confidence interval

coverage at the tail. Interestingly, the jackknife is quite competitive with the proposed method in terms of coverage performance despite the poor performance of the MSE.

Table 4 shows the results under the model $U^* \sim$ Unif(−5, 15), $T^* \sim$ Unif(0, 15), and $V^* = U^* + 5$. All the three variance estimators are nearly unbiased and their MSEs are very similar. Although the bootstrap seems to provide the best result in terms of the MSE, the three methods are quite competitive. In terms of coverage probability, the bootstrap tends to be the best.

Although we found no single best method across all criteria, the advantage of the proposed method over other methods appears for larger samples (n = 250 and 300). The MSE of the proposed method is smallest in majority of cases. Unlike the bootstrap that may exhibit serious under-coverage at the tails, the proposed method can alleviate the problem for large sample sizes. As for the computational cost among the three methods, the proposed method is the lowest since it merely performs the matrix algebra in Eq. (4). On the other extreme, the bootstrap requires performing the

Table 3 Simulation results under $U^* \sim \text{Unif}(0,\ a)$, $T^* \sim \text{Unif}(0,\ 1)$, and $V^* \sim \text{Unif}(b,\ 1)$ with $a = 0.5$ and $b = 0.5$ based on 500 replications

		$n = 100$	$n = 150$	$n = 200$	$n = 250$	$n = 300$
$F(t) = 0.5$						
SD		0.093	0.078	0.068	0.059	0.055
ESD	Proposed	0.084	0.070	0.060	0.054	0.049
	Bootstrap	0.086	0.071	0.061	0.055	0.050
	Jackknife	0.092	0.076	0.063	0.057	0.051
MSE	Proposed	0.00237	0.00227	0.00141	0.00069	0.00059
	Bootstrap	0.00125	0.00104	0.00061	0.00054	0.00040
	Jackknife	0.00385	0.00326	0.00161	0.00150	0.00087
95%Cov	Proposed	0.932	0.934	0.950	0.952	0.958
	Bootstrap	0.934	0.944	0.954	0.952	0.958
	Jackknife	0.934	0.944	0.950	0.964	0.962
$F(t) = 0.2$						
SD		0.093	0.074	0.065	0.059	0.052
ESD	Proposed	0.080	0.066	0.058	0.052	0.048
	Bootstrap	0.084	0.069	0.060	0.054	0.049
	Jackknife	0.090	0.073	0.061	0.055	0.050
MSE	Proposed	0.00270	0.00132	0.00077	0.00055	0.00035
	Bootstrap	0.00237	0.00152	0.00087	0.00071	0.00046
	Jackknife	0.00551	0.00366	0.00138	0.00119	0.00051
95%Cov	Proposed	0.932	0.938	0.946	0.928	0.926
	Bootstrap	0.908	0.902	0.934	0.924	0.910
	Jackknife	0.944	0.950	0.952	0.942	0.938

$\text{ESD} = \frac{1}{500}\sum_{r=1}^{500}\sqrt{\hat{V}\{\hat{F}(t)\}_{(r)}}$

$\text{MSE} = \frac{1}{500}\sum_{r=1}^{500}\left(\sqrt{\hat{V}\{\hat{F}(t)\}_{(r)}} - \text{SD}\{\hat{F}(t)\}\right)^2$

95%Cov = Empirical coverage probability of the 95% confidence interval

self-consistency algorithms over $B = 1{,}000$ resamplings. Hence, the proposed method would be useful when the sample size is large.

5.3 Performance of the goodness-of-fit test

First, we examine the type I error of the goodness-of-fit tests introduced in Sect. 4.2. For each run, we record the rejection/acceptance status of the goodness-of-fit tests at the $100\alpha\%$ level, and calculate the rejection rates among 500 repetitions. We also compare the null means of the tests (denoted by $E[C]$ and $E[K]$) with the resampling means (denoted by $E[C^{(b)}]$ and $E[K^{(b)}]$).

As shown in Table 5, the rejection rates (type I error rates) are in good agreement with the selected nominal sizes ($\alpha = 0.01, 0.05$, and 0.10). In addition, the sample means $E[C]$ and $E[K]$ are close to the resampling means $E[C^{(b)}]$ and $E[K^{(b)}]$, respectively.

Table 4 Simulation results under $U^* \sim$ Unif(−5, 15), $T^* \sim$ Unif(0, 15), and $V^* = U^* + 5$ based on 500 replications

		$n = 100$	$n = 150$	$n = 200$	$n = 250$	$n = 300$
$F(t) = 0.5$						
SD		0.146	0.121	0.103	0.096	0.086
ESD	Proposed	0.146	0.120	0.105	0.094	0.086
	Bootstrap	0.147	0.120	0.105	0.094	0.086
	Jackknife	0.156	0.125	0.108	0.096	0.088
MSE	Proposed	0.00064	0.00023	0.00013	0.000065	0.000046
	Bootstrap	0.00048	0.00020	0.00010	0.000069	0.000046
	Jackknife	0.00089	0.00026	0.00014	0.000067	0.000047
95%Cov	Proposed	0.904	0.932	0.946	0.928	0.928
	Bootstrap	0.940	0.944	0.950	0.950	0.930
	Jackknife	0.910	0.936	0.946	0.940	0.936
$F(t) = 0.2$						
SD		0.101	0.084	0.071	0.064	0.057
ESD	Proposed	0.100	0.081	0.070	0.062	0.057
	Bootstrap	0.106	0.083	0.072	0.063	0.058
	Jackknife	0.107	0.084	0.072	0.064	0.058
MSE	Proposed	0.00109	0.00048	0.00027	0.00018	0.00012
	Bootstrap	0.00097	0.00044	0.00024	0.00017	0.00012
	Jackknife	0.00133	0.00053	0.00028	0.00018	0.00013
95%Cov	Proposed	0.938	0.944	0.946	0.942	0.948
	Bootstrap	0.958	0.946	0.954	0.944	0.948
	Jackknife	0.956	0.954	0.952	0.942	0.954

ESD = $\frac{1}{500}\sum_{r=1}^{500}\sqrt{\hat{V}\{\hat{F}(t)\}_{(r)}}$

MSE = $\frac{1}{500}\sum_{r=1}^{500}\left(\sqrt{\hat{V}\{\hat{F}(t)\}_{(r)}} - \mathrm{SD}\{\hat{F}(t)\}\right)^2$

95%Cov = Empirical coverage probability of the 95 % confidence interval

However, under $(a, b) = (0.67, 0.33)$, the Cramér–von Mises test leads to somewhat higher rejection rates than the nominal sizes. Overall, the Kolmogorov–Smirnov test produces a slightly conservative result.

Next, we examine the power under alternative hypotheses. We focus on the case of $(a, b) = (0.5, 0.5)$ under the null $F_0(t) = t\mathbf{I}(0 < t < 1)$ and alternatives

(1) $F_1(t) = t^{1/\gamma}\mathbf{I}(0 < t < 1)$, γ=1/1.8, 1/1.6, ..., 1, ..., 1.6, 1.8.
(2) $F_2(t) = t\mathbf{I}(0 < t < \gamma)/\gamma$, γ=1, 0.975, 0.95, ..., 0.75, 0.725.

As shown in Fig. 2, the power increases as γ departs from the null model of $\gamma = 1$. The curves for $\alpha = 0.05$ (right panels) are consistently higher than those for $\alpha = 0.01$ (left panels). It is found that the Cramér–von Mises test exhibits higher power than the Kolmogorov–Smirnov test under the alternative model (1). This conclusion, however, should not be overemphasized as the type I error rates of the Cramér–von Mises test

Table 5 Simulation results for the proposed goodness-of-fit tests under the null hypothesis based on 500 replications

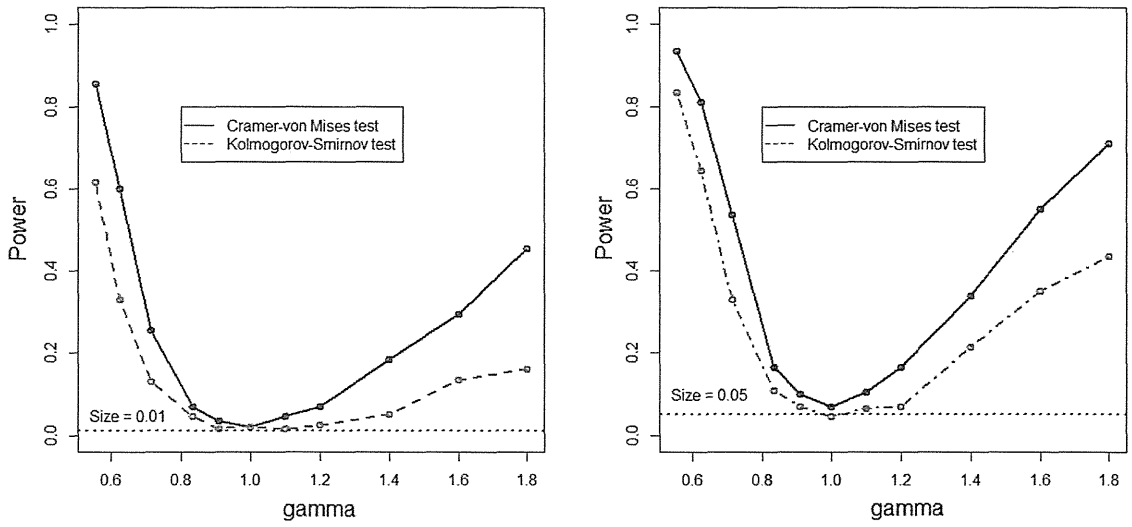
		Cramér–von Mises test (C)			Kolmogorov–Smirnov test (K)		
		n			n		
		100	150	250	100	150	250
$(a,\ b) =$ (0.25, 0.75)	Reject rate at $\alpha = 0.10$	0.096	0.097	0.100	0.075	0.075	0.095
	Reject rate at $\alpha = 0.05$	0.045	0.047	0.056	0.031	0.037	0.043
	Reject rate at $\alpha = 0.01$	0.006	0.011	0.018	0.006	0.005	0.015
	$E[C]$ or $E[K]$	0.608	0.623	0.667	0.118	0.099	0.081
	$E[C^{(b)}]$ or $E[K^{(b)}]$	0.625	0.614	0.605	0.115	0.096	0.077
$(a,\ b) =$ (0.5, 0.5)	Reject rate at $\alpha = 0.10$	0.120	0.105	0.088	0.089	0.081	0.078
	Reject rate at $\alpha = 0.05$	0.063	0.055	0.045	0.037	0.033	0.030
	Reject rate at $\alpha = 0.01$	0.015	0.014	0.008	0.006	0.007	0.005
	$E[C]$ or $E[K]$	1.078	1.206	1.306	0.143	0.121	0.098
	$E[C^{(b)}]$ or $E[K^{(b)}]$	1.167	1.281	1.331	0.137	0.116	0.093
$(a,\ b) =$ (0.67, 0.33)	Reject rate at $\alpha = 0.10$	0.140	0.150	0.120	0.090	0.065	0.105
	Reject rate at $\alpha = 0.05$	0.095	0.085	0.060	0.040	0.035	0.040
	Reject rate at $\alpha = 0.01$	0.015	0.020	0.030	0.010	0.005	0.010
	$E[C]$ or $E[K]$	1.006	1.109	0.989	0.142	0.118	0.091
	$E[C^{(b)}]$ or $E[K^{(b)}]$	1.001	0.984	0.820	0.135	0.113	0.087
$c = 5$	Reject rate at $\alpha = 0.10$	0.108	0.119	0.119	0.096	0.106	0.110
	Reject rate at $\alpha = 0.05$	0.063	0.057	0.059	0.052	0.055	0.055
	Reject rate at $\alpha = 0.01$	0.015	0.014	0.014	0.010	0.011	0.013
	$E[C]$ or $E[K]$	0.417	0.411	0.412	0.104	0.084	0.067
	$E[C^{(b)}]$ or $E[K^{(b)}]$	0.403	0.403	0.401	0.103	0.085	0.066

The average of the Cramér–von Mises statistics is denoted by $E[C]$. The average of its resampling version is denoted by $E[C^{(b)}]$. $E[K]$ and $E[K^{(b)}]$ are defined similarly for the Kolmogorov–Smirnov statistics

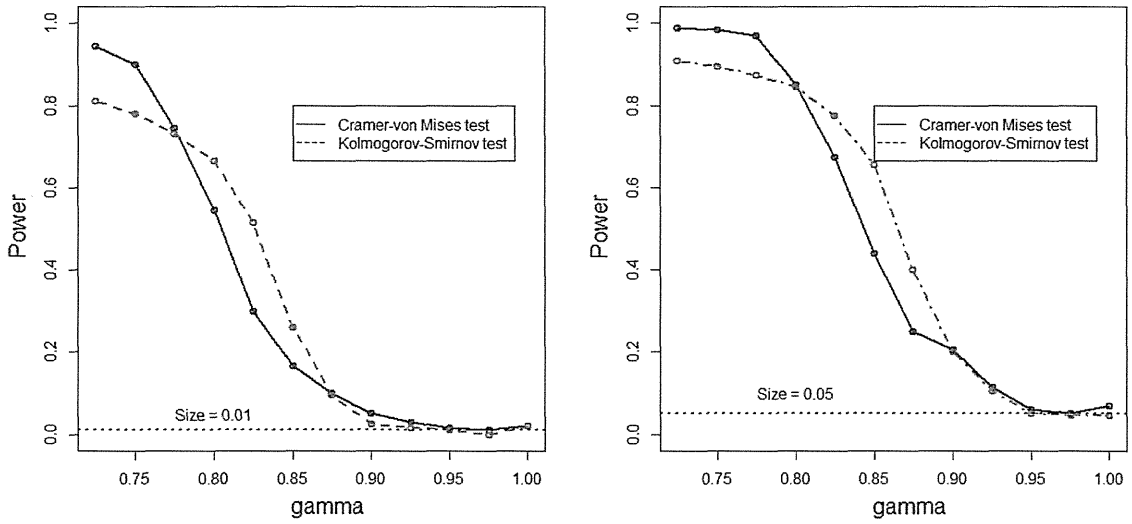
are slightly higher than those of the Kolmogorov–Smirnov test. The results for other $(a,\ b)$ are similar.

5.4 Performance of the confidence band

We investigate the coverage performance of the EP and HW bands introduced in Sect. 4.3. The EP band is calculated under p_1 =0.2 and p_2 =0.8. For each run, we check if the bands completely cover the true F. The coverage rates over the 500 replications are given in Table 6. Overall, the coverage rates reflect the nominal levels and are particularly accurate when $1 - \alpha = 0.99$. The EP band has slightly more accurate coverage compared to the HW band, especially at levels $1 - \alpha = 0.90$ and 0.95. This is because the HW band exhibits slight over-coverage, which parallels the conservative results of the Kolmogorov–Smirnov test.



(1) Alternative $F_1(t) = t^{1/\gamma}\mathbf{I}(0 < t < 1)$, γ=1/1.8, 1/1.6, ..., 1, ..., 1.6, 1.8.



(2) Alternative $F_2(t) = t\mathbf{I}(0 < t < \gamma)/\gamma$, γ=1, 0.975, 0.95, ..., 0.75, 0.725

Fig. 2 The power curves for the proposed goodness-of-fit tests with sizes α =0.01 (*left panel*) and α =0.05 (*right panel*) based on n = 150. The value $\gamma = 1$ corresponds to the null, while $\gamma \neq 1$ corresponds to the alternative. (1) Alternative $F_1(t) = t^{1/\gamma}\mathbf{I}(0 < t < 1)$, γ=1/1.8, 1/1.6, ..., 1, ..., 1.6, 1.8. (2) Alternative $F_2(t) = t\mathbf{I}(0 < t < \gamma)/\gamma$, γ=1, 0.975, 0.95, ..., 0.75, 0.725

6 Data analysis

We analyzed the childhood cancer data from Moreira and Uña-Álvarez (2010) as described in Sect. 1. The sample consists of 409 children with $\{ (U_j, T_j, V_j) : j = 1, \ldots, 409 \}$ subject to double-truncation $U_j \leq T_j \leq V_j$, where T_j is the age (in days) at diagnosis, U_j is the age at the start of follow-up (January 1, 1999), and $V_i = U_i + 1825$ is the age at the end of follow-up (December 31, 2003). The primary interest here is inference of the distribution function $F(t) = \Pr(T^* \leq t)$, where T^* is the pre-truncated age at diagnosis. We depict the NPMLE $\hat{F}(t)$ in Fig. 3. The resulting curve is virtually identical to that reported in Moreira and Uña-Álvarez (2010). They

Table 6 Coverage rates of the proposed confidence bands at the $100(1-\alpha)\%$ level based on 500 replications

	Nominal level	$n = 100$	$n = 150$	$n = 250$
EP (equal precision) band				
$a = 0.25, b = 0.75$	$1-\alpha = 0.900$	0.904	0.924	0.902
	$1-\alpha = 0.950$	0.958	0.952	0.950
	$1-\alpha = 0.990$	0.990	0.990	0.984
$a = 0.5, b = 0.5$	$1-\alpha = 0.900$	0.908	0.910	0.918
	$1-\alpha = 0.950$	0.964	0.954	0.960
	$1-\alpha = 0.990$	0.990	0.988	0.990
$a = 0.67, b = 0.33$	$1-\alpha = 0.900$	0.915	0.905	0.910
	$1-\alpha = 0.950$	0.950	0.955	0.950
	$1-\alpha = 0.990$	0.985	0.995	0.985
$c = 5$	$1-\alpha = 0.900$	0.894	0.894	0.876
	$1-\alpha = 0.950$	0.928	0.940	0.932
	$1-\alpha = 0.990$	0.984	0.986	0.986
HW (Hall–Wellner) band				
$a = 0.25, b = 0.75$	$1-\alpha = 0.900$	0.927	0.926	0.905
	$1-\alpha = 0.950$	0.969	0.963	0.957
	$1-\alpha = 0.990$	0.994	0.995	0.985
$a = 0.5, b = 0.5$	$1-\alpha = 0.900$	0.912	0.919	0.922
	$1-\alpha = 0.950$	0.963	0.967	0.970
	$1-\alpha = 0.990$	0.994	0.993	0.995
$a = 0.67, b = 0.33$	$1-\alpha = 0.900$	0.910	0.935	0.895
	$1-\alpha = 0.950$	0.960	0.965	0.960
	$1-\alpha = 0.990$	0.990	0.995	0.990
$c = 5$	$1-\alpha = 0.900$	0.904	0.894	0.890
	$1-\alpha = 0.950$	0.948	0.945	0.945
	$1-\alpha = 0.990$	0.990	0.989	0.987

provide pointwise confidence intervals using the bootstrap. In this paper, we provide additional inference procedures using goodness-of-fit tests and confidence bands.

For goodness-of-fit tests, we set the following two hypotheses:

$$H_{01}: \quad F(t) = \frac{t}{5475}\mathbf{I}(0 < t < 5475) + \mathbf{I}(t \geq 5475)$$

and

$$H_{02}: \quad F(t) = \left(\frac{t}{5475}\right)^{3/4}\mathbf{I}(0 < t < 5475) + \mathbf{I}(t \geq 5475),$$

where $5{,}475 = 15 \times 365$ (days) is the maximum age to be defined as childhood cancer (15 years old). Here, H_{01} implies that childhood cancer occurs uniformly over all ages

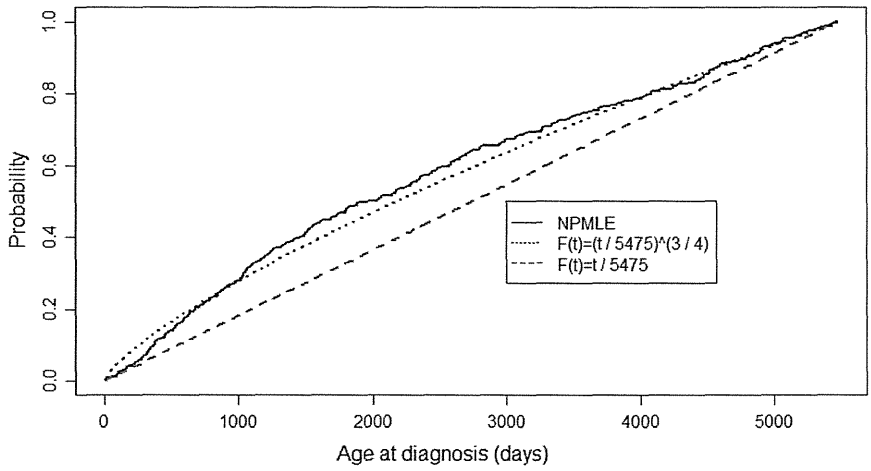


Fig. 3 The NPMLE $\hat{F}(t)$ of the distribution function of ages at diagnosis for childhood cancer (*solid line*). The hypothesized curves are $H_{01}: F(t) = (t/5475)\ \mathbf{I}(0 < t < 5475) + \mathbf{I}(t \geq 5475)$ (*dashed line*), $H_{02}: F(t) = (t/5475)^{3/4}\mathbf{I}(0 < t < 5475) + \mathbf{I}(t \geq 5475)$ (*dotted line*)

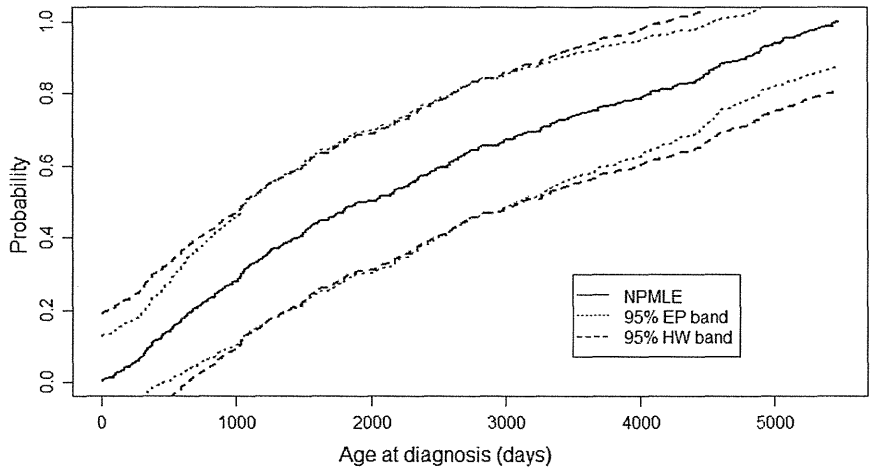


Fig. 4 The NPMLE and its 95% confidence bands. The *dotted line* is the EP (equal precision) band and the *dashed line* is the HW (Hall–Wellner) band

under 15 years, while H_{02} implies that the occurrence of childhood cancer decreases as their age increases. Figure 3 depicts the two hypothesized curves along with the NPMLE. The curve for H_{02} fits better than the curve for H_{01}. Indeed, the Cramér–von Mises test rejects H_{01} at 10 % significance level (P-value = 0.094), while does not reject H_{02} (P-value = 0.732). Similar results are found through the Kolmogorov–Smirnov test (P-value = 0.099 for H_{01} and P-value = 0.797 for H_{02}).

Figure 4 displays the 95 % EP and HW bands based on the algorithm in Sect. 4.3. The EP band is calculated under p_1 =0.1 and p_2 =0.9. The EP and HW bands are generally competitive, but the EP band is slightly narrower in the tails. This is qualitatively similar to the EP and HW bands for right-censored data. Now, we validate the coverage performance using the bootstrap as mentioned in Sect. 4.3. The EP band covers 950 out of the 1000 bootstrap NPMLEs and the HW band covers 964 out of the 1000 bootstrap NPMLEs. Hence, the coverage level is close to the nominal 95 %.

We compare the three variance estimators (proposed, bootstrap and jackknife) for selected values of t. The computation time required for the three estimators are also

Table 7 Variance estimates of the NPMLE based on the childhood cancer data

	Proposed: $\sqrt{\hat{V}_{\text{Info}}\{\hat{F}(t)\}}$	Bootstrap: $\sqrt{\hat{V}_{\text{Boot}}\{\hat{F}(t)\}}$	Jackknife: $\sqrt{\hat{V}_{\text{Jack}}\{\hat{F}(t)\}}$
Variance estimate at $t = 750.0$	0.0469	0.0464	0.0473
Computation time (s)	(0.25)	(342.16)	(118.37)
Variance estimate at $t = 2{,}083.5$	0.0817	0.0814	0.0815
Computation time (s)	(0.22)	(313.53)	(115.83)
Variance estimate at $t = 4{,}251.0$	0.0599	0.0665	0.0644
Computation time (s)	(0.23)	(317.48)	(117.37)

The three variance estimators are calculated at $t = 750.0$, $t = 2{,}083.5$, and $t = 4{,}251.0$, corresponding to the 20, 50, and 80 percentiles of observed ages at diagnosis, respectively. Required computation times for the three methods are also compared

compared. As shown in Table 7, the three estimates produce very similar results for all t. On the other hand, the computation time required for the proposed method is much shorter than those of the resampling-based methods.

7 Conclusion

This paper introduced a simple and explicit covariance estimator of the NPMLE using the observed information matrix. This technique provides various inference procedures, including pointwise confidence interval, goodness-of-fit, and confidence band.

Our simulations showed that the major advantage of the proposed variance estimator over the bootstrap and jackknife was for the larger samples ($n = 250$ and 300). The data analysis demonstrated the reduced computational time for the proposed method vis-à-vis the bootstrap and jackknife methods. Hence, the proposed method is most useful when the sample size is very large, which often occurs in demography and epidemiology (e.g., Stovring and Wang 2007). In such large-scale studies, the proposed method may be the best possible choice for statistical inference.

For goodness-of-fit procedures, we developed the Kolmogorov–Smirnov and Cramér–von Mises tests with the null distributions simulated by the proposed covariance structure. The simulations showed that these tests have proper type I error rates and power. Applying the tests to the childhood cancer data, we rejected the scientific assumption that childhood cancer occurs uniformly over all ages below 15 years. This conclusion could not have been derived without developing the goodness-of-fit procedures.

Acknowledgments We would like to thank the editor, the associate editor and the two reviewers for their helpful comments and corrections that greatly improved the manuscript. This work was financially supported by the National Science Council of Taiwan (NSC101-2118-M008-002-MY2) to T. Emura, and a Grant-in-Aid for a Research Fellow of the Japan Society for the Promotion of Science to H. Michimae (No. 23570036). The work of Y. Konno was partially supported by Grant-in-Aid for Scientific Research(C) (No. 25330043 and 21500283).

Appendix A: Bootstrap and jackknife algorithms

Simple bootstrap algorithm (Moreira and Uña-Álvarez 2010):

Step 1: For each $b = 1, \ldots, B$, draw bootstrap resamples $\{ (U_{jb}^*, T_{jb}^*, V_{jb}^*) : j = 1, \ldots, n \}$ from $\{ (U_j, T_j, V_j) : j = 1, \ldots, n \}$, and then compute the NPMLE $\hat{F}_b^*(t)$ from them.
Step 2: Compute the bootstrap variance estimator

$$\hat{V}_{\text{Boot}}\{\hat{F}(t)\} = \frac{1}{B-1}\sum_{b=1}^{B}\{\hat{F}_b^*(t) - \bar{F}^*(t)\}^2,$$

where $\bar{F}^*(t) = \frac{1}{B}\sum_{b=1}^{B}\hat{F}_b^*(t)$, and take the $(\alpha/2)\times 100\%$ and $(1-\alpha/2)\times 100\%$ points of $\{ \hat{F}_b^*(t) : b = 1, \ldots, B \}$ for the$(1-\alpha)\times 100\%$ confidence interval.

Jackknife algorithm:

Step 1: For each $i = 1, \ldots, n$, delete the i th sample from $\{ (U_j, T_j, V_j) : j = 1, \ldots, n \}$, and then compute the NPMLE $\hat{F}_{(-i)}(t)$ from the remaining $n-1$ samples.
Step 2: Compute the jackknife variance estimator

$$\hat{V}_{\text{Jack}}\{\hat{F}(t)\} = \frac{n-1}{n}\sum_{i=1}^{n}\{\hat{F}_{(-i)}(t) - \bar{F}_{(\cdot)}(t)\}^2,$$

where $\bar{F}_{(\cdot)}(t) = \frac{1}{n}\sum_{i=1}^{n}\hat{F}_{(-i)}(t)$, and the log-transformed $(1-\alpha)\times 100\%$ confidence interval

$$\left(\hat{F}(t)\exp[\, -z_{\alpha/2}\hat{V}_{\text{Jack}}^{1/2}\{\hat{F}(t)\}/\hat{F}(t)\,],\quad \hat{F}(t)\exp[\, z_{\alpha/2}\hat{V}_{\text{Jack}}^{1/2}\{\hat{F}(t)\}/\hat{F}(t)\,] \right).$$

Appendix B: Asymptotic theory

Appendix B1. Weak convergence of $\sqrt{n}(\,\hat{F}(t) - F(t)\,)$

Although not stated explicitly, we assume that the identifiability conditions (Shen 2010, p. 836) are satisfied. Consider the log-likelihood function

$$\ell_n(F)/n = \sum_{i=1}^{n}(\log f_j - \log F_j)/n.$$

For any $h \in Q$, where Q is the set of all uniformly bounded functions, let $H(t) = \int_0^t h(s)dF(s)$ and $\hat{H}(t) = \int_0^t h(s)d\hat{F}(s)$ where h satisfies the constraint $\hat{H}(\infty) = 1$. Suppose that $\hat{F}$ is the maximizer of $\ell_n(F)$. Then for any $h \in Q$ and $\varepsilon \geq 0$, we have $\ell_n(\hat{F} + \varepsilon\hat{H}) \leq \ell_n(\hat{F})$. Hence, the score function $\partial\ell_n(F + \varepsilon H)/\partial\varepsilon|_{\varepsilon=0}$ is equal to

$$\Psi_n(F)[h] \equiv \frac{1}{n}\sum_{i=1}^{n}\left[h(T_i) - \frac{\int \mathbf{I}(U_i \le s \le V_i)h(s)dF(s)}{\int \mathbf{I}(U_i \le s \le V_i)dF(s)}\right],$$

for any $h \in Q$. The expectation is defined as

$$\Psi(F)[h] \equiv E\left[h(T^*) - \frac{\int \mathbf{I}(U^* \le s \le V^*)h(s)dF(s)}{\int \mathbf{I}(U^* \le s \le V^*)dF(s)}\right].$$

Consider $\Psi_n(F)[h]$ as a random function defined on Q. Accordingly, consider a random map $\Theta \to l^\infty(Q)$, defined by $F \mapsto \Psi_n(F)[\cdot]$. Then, the equation $\Psi_n(F)[\cdot] = 0$ is considered the estimating function that takes its value on $l^\infty(Q)$. It follows that the NPMLE is the Z-estimator that satisfies $\Psi_n(\hat{F})[\cdot] = 0$ (van der Vaart and Wellner 1996, p. 309). In the following, we assume that certain regularity conditions for the asymptotic theory for the Z-estimator hold, which include the asymptotic approximation condition, the Fréchet differentiability of the map, and the invertibility of the derivative map.

Then, one can write

$$0 = n^{1/2}\Psi_n(\hat{F})[h] = n^{1/2}\Psi_n(F)[h] + n^{1/2}\dot{\Psi}_F(\hat{F} - F)[h] + o_P(1), \tag{5}$$

where $\dot{\Psi}_F(\hat{F} - F)[h]$ is the derivative of $\Psi_n(F)[h]$ at F with direction $\hat{F} - F$. It follows from the form of $\Psi(F)[\cdot]$ that

$$\dot{\Psi}_F(\hat{F} - F)[h] = \frac{d}{dt}\Psi\{\ \hat{F} + t(\hat{F} - F)\ \}[h]|_{t=0} = -\int \sigma_F(x)[h]d(\hat{F} - F)(x). \tag{6}$$

It follows from Eqs. (5) and (6) that the NPMLE satisfies the asymptotic linear expression

$$\begin{aligned}&\sqrt{n}\int \sigma_F(x)[h]d(\hat{F} - F)(x)\\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[h(T_i) - \frac{\int I(U_i \le s \le V_i)h(s)dF(s)}{\int I(U_i \le s \le V_i)dF(s)}\right] + o_P(1),\end{aligned} \tag{7}$$

where the right-side converges weakly to a mean zero Gaussian process with the covariance structure

$$\begin{aligned}&E\left[h(T^*) - \frac{\int \mathbf{I}(U^* \le s \le V^*)h(s)dF(s)}{\int \mathbf{I}(U^* \le s \le V^*)dF(s)}\right]\left[h'(T^*) - \frac{\int \mathbf{I}(U^* \le s \le V^*)h'(s)dF(s)}{\int \mathbf{I}(U^* \le s \le V^*)dF(s)}\right]\\ &= \int \sigma_F(x)[h]h'(x)dF(x),\end{aligned}$$

for bounded functions h and h'. The desired weak convergence of $\sqrt{n}(\ \hat{F}(t) - F(t)\ )$ is obtained by setting $h = \sigma_F^{-1}(w_t)$ in Eq. (7).

Appendix B2: Proof of $\sum_{j=1}^{n} w_s(T_j)\hat{\sigma}_F^{-1}(w_t)(T_j)\hat{f}_j = \mathbf{W}_s^{\mathrm{T}} \left\{ \frac{i_n(\mathbf{f})}{n} \right\}^{-1} \mathbf{W}_t$

It follows that

$$\hat{\sigma}_F(T_j)[h] = \frac{1}{n}\sum_{i=1}^{n} J_{ij}\left\{\frac{h_j}{\hat{F}_i} - \frac{1}{\hat{F}_i^2}\sum_{k=1}^{n} J_{ik}h_k\hat{f}_k\right\} = \frac{1}{n}\left[\frac{h_j\hat{f}_j}{\hat{f}_j^2} - \sum_{i=1}^{n}\sum_{k=1}^{n}\frac{J_{ij}J_{ik}}{\hat{F}_i^2}h_k\hat{f}_k\right]. \tag{8}$$

Note that

$$J^{\mathrm{T}}\mathrm{diag}\left(\frac{1}{\mathbf{F}^2}\right)J = \begin{bmatrix} \sum_{i=1}^{n}\frac{J_{i1}J_{i1}}{F_i^2} & \cdots & \sum_{i=1}^{n}\frac{J_{i1}J_{in}}{F_i^2} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{n}\frac{J_{in}J_{i1}}{F_i^2} & \cdots & \sum_{i=1}^{n}\frac{J_{in}J_{in}}{F_i^2} \end{bmatrix}.$$

Hence, Eq. (8) with $h = h'$ and $\sigma_F(x)[h'] = w_t(x) = \mathbf{I}(x \le t)$ yield

$$\begin{aligned} \begin{bmatrix} w_t(T_1) \\ \vdots \\ w_t(T_n) \end{bmatrix} &= \frac{1}{n}\left[\left\{\mathrm{diag}\left(\frac{1}{\hat{\mathbf{f}}^2}\right) - J^{\mathrm{T}}\mathrm{diag}\left(\frac{1}{\hat{\mathbf{F}}^2}\right)J\right\}\Big|_{\hat{f}_n = 1 - \mathbf{1}_{n-1}^{\mathrm{T}}\hat{\mathbf{f}}}\right]\begin{bmatrix} h_1\hat{f}_1 \\ \vdots \\ h_n\hat{f}_n \end{bmatrix} \\ &= \frac{1}{n}\left[\left\{\mathrm{diag}\left(\frac{1}{\hat{\mathbf{f}}^2}\right) - J^{\mathrm{T}}\mathrm{diag}\left(\frac{1}{\hat{\mathbf{F}}^2}\right)J\right\}\Big|_{\hat{f}_n = 1 - \mathbf{1}_{n-1}^{\mathrm{T}}\hat{\mathbf{f}}}\right] D^{\mathrm{T}}\begin{bmatrix} h_1\hat{f}_1 \\ \vdots \\ h_{n-1}\hat{f}_{n-1} \end{bmatrix}, \end{aligned}$$

where the last equation uses the constraint $\sum_{j=1}^{n} h_j\hat{f}_j = 0$. Multiplying D for both sides, and taking the inverse of the information matrix,

$$\begin{bmatrix} \hat{\sigma}_F^{-1}(w_t)(T_1)\hat{f}_1 \\ \vdots \\ \hat{\sigma}_F^{-1}(w_t)(T_{n-1})\hat{f}_{n-1} \end{bmatrix} = \left\{\frac{i_n(\hat{\mathbf{f}})}{n}\right\}^{-1}\begin{bmatrix} w_t(T_1) - w_t(T_n) \\ \vdots \\ w_t(T_1) - w_t(T_n) \end{bmatrix}.$$

It follows that

$$\begin{aligned} \sum_{j=1}^{n} w_s(T_j)\hat{\sigma}_F^{-1}(w_t)(T_j)\hat{f}_j &= \sum_{j=1}^{n-1}\{\, w_s(T_j) - w_s(T_n)\,\}\hat{\sigma}_F^{-1}(w_t)(T_j)\hat{f}_j \\ &= \left[\, w_s(T_1) - w_s(T_n) \cdots w_s(T_{n-1}) - w_s(T_n)\,\right]\left\{\frac{i_n(\hat{\mathbf{f}})}{n}\right\}^{-1}\begin{bmatrix} w_t(T_1) - w_t(T_n) \\ \vdots \\ w_t(T_1) - w_t(T_n) \end{bmatrix} \\ &= \mathbf{W}_s^{\mathrm{T}}\left\{\frac{i_n(\hat{\mathbf{f}})}{n}\right\}^{-1}\mathbf{W}_t. \end{aligned}$$

References

Austin D, Simon DK, Betensky RA (2013) Computationally simple estimation and improved efficiency for special cases of double truncation. Lifetime Data Anal. doi:10.1007/s10985-013-9287-z

Chen YH (2010) Semiparametric marginal regression analysis for dependent competing risks under an assumed copula. J R Stat Soc B 72:235–251

Commenges D (2002) Inference for multi-state models from interval-censored data. Stat Methods Med Res 11:167–182

Efron B, Petrosian V (1999) Nonparametric method for doubly truncated data. J Am Stat Assoc 94:824–834

Emura T, Wang W (2012) Nonparametric maximum likelihood estimation for dependent truncation data based on copulas. J Multivar Anal 110:171–188

Emura T, Konno Y (2012) Multivariate normal distribution approaches for dependently truncated data. Stat Papers 53:133–149

Klein JP, Moeschberger ML (2003) Survival analysis: techniques for censored and truncated data. Springer, New York

Moreira C, Uña-Álvarez J (2010) Bootstrapping the NPMLE for doubly truncated data. J Nonparametr Stat 22:567–583

Moreira C, Uña-Álvarez J (2012) Kernel density estimation with doubly-truncated data. Electron J Stat 6:501–521

Moreira C, Keilegom IV (2013) Bandwidth selection for kernel density estimation with doubly truncated data. Comput Stat Data Anal 61:107–123

Moreira C, Uña-Álvarez J, Meira-Machado L (2014) Nonparametric regression with doubly truncated data. Comput Stat Data Anal. doi:10.1016/j.csda.2014.03.017

Murphy SA (1995) Asymptotic theory for the frailty model. Ann Stat 23:182–198

Nair VN (1984) Confidence bands for survival functions with censored data: a comparative study. Technometrics 26:265–275

Shen PS (2010) Nonparametric analysis of doubly truncated data. Ann Inst Stat Math 62:835–853

Shen PS (2011) Testing quasi-independence for doubly truncated data. J Nonparametr Stat 23:1–9

Shen PS (2012) Empirical likelihood ratio with doubly truncated data. J Appl Stat 38:2345–2353

Stovring H, Wang MC (2007) A new approach of nonparametric estimation of incidence and lifetime risk based on birth rates and incidence events. BMC Med Res Methodol 7:53

van der Vaart AW, Wellner JA (1996) Weak convergence and empirical process. Springer-Varlag, New York

Zeng D, Lin DY (2006) Efficient estimation of semiparametric transformation models for counting processes. Biometrika 93:627–640

Zhu H, Wang MC (2012) Analyzing bivariate survival data with interval sampling and application to cancer epidemiology. Biometrika 99:345–361

Gynecologic Cancer InterGroup (GCIG) Consensus Review for Clear Cell Carcinoma of the Ovary

Aikou Okamoto, MD, PhD, Rosalind M. Glasspool, MBBS, PhD, FRCP,† Seiji Mabuchi, MD, PhD,‡ Noriomi Matsumura, MD, PhD,§ Hiroyuki Nomura, MD, PhD,|| Hiroaki Itamochi, MD, PhD,¶ Masashi Takano, MD, PhD,# Tadao Takano, MD, PhD,** Nobuyuki Susumu, MD, PhD,|| Daisuke Aoki, MD, PhD,|| Ikuo Konishi, MD, PhD,§ Alan Covens, MD, LMCC, FRCSC,†† Jonathan Ledermann, MD, FRCP,‡‡ Delia Mezzazanica, PhD,§§ Christopher Steer, MBBS, FRACP,|||| David Millan, BSc, MB, ChB, FRCPath,¶¶ Iain A. McNeish, MD, PhD,## Jacobus Pfisterer, MD, PhD,*** Sokbom Kang, MD, PhD,††† Laurence Gladieff, MD,‡‡‡ Jane Bryce, MSN,§§§ and Amit Oza, MD, FRPCPC, MBBs||||||*

Abstract: Clear cell carcinoma of the ovary (CCC) is a histologic subtype of epithelial ovarian cancer with a distinct clinical behavior. There are marked geographic differences in the prevalence of CCC. The CCC is more likely to be detected at an early stage than high-grade serous cancers, and when confined within the ovary, the prognosis is good. However, advanced disease is associated with a very poor prognosis and resistance to standard treatment. Cytoreductive surgery should be performed for patients with stage II, III, or IV disease. An international phase III study to compare irinotecan/cisplatin and paclitaxel/carboplatin as adjuvant chemotherapy for stage IIV CCC has completed enrollment (GCIG/JGOG3017). Considering the frequent *PIK3CA* mutation in CCC, dual inhibitors targeting PI3K, AKT in the mTOR pathway, are promising. Performing these trials and generating the evidence will require considerable international collaboration.

Key Words: Clear cell carcinoma of the ovary (CCC), Deep venous thrombosis, Glycogen, Hepatocyte nuclear factor-1ß, *WT1*, Pulmonary embolism, Ethnics, *ARID1A*, *PIK3CA*, *PPM1D*, *PPP2R1A*, *KRAS*, Cytoreductive surgery, Paclitaxel, Platinum, Irinotecan hydrochloride, Cisplatin, PI3K/AKT/mTOR

Received May 27, 2014, and in revised form August 26, 2014.
Accepted for publication August 31, 2014.

(*Int J Gynecol Cancer* 2014;24: S20–S25)

*Jikei University School of Medicine, Tokyo, Japan; †Beatson West of Scotland Cancer Centre, Glasgow, United Kingdom; ‡Osaka University Graduate School of Medicine, Osaka, Japan; §Kyoto University Graduate School of Medicine, Kyoto, Japan; ||Keio University, Tokyo, Japan; ¶Tottori University School of Medicine, Tottori, Japan; #National Defense Medical College, Saitama, Japan; **Clinical Research, Innovation, and Education Center, Tohoku University Hospital, Sendai, Japan; ††Sunnybrook Hospital, Toronto, Canada; ‡‡UCL Cancer Institute, London, United Kingdom; §§Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy; ||||Border Medical Oncology, Victoria, Australia; ¶¶Department of Pathology, Glasgow Royal Infirmary, Glasgow, United Kingdom; ##Institute of Cancer Sciences, University of Glasgow Wolfson Wohl Cancer Research Centre, Beatson Institute for Cancer Research, Glasgow, United Kingdom; ***Gynecologic Oncology Center, Kiel, Germany; †††National Cancer Center, Goyang, Korea; ‡‡‡Institut Claudius Regaud, Toulouse, France; §§§National Cancer Institute, Naples, Italy; and ||||||Princess Margaret Hospital, Toronto, Canada.

Address correspondence and reprint requests to Aikou Okamoto, MD, PhD, Department of Obstetrics and Gynecology, The Jikei University School of Medicine, 3-25-8 Nishi-Shinbashi, Minato-ku Tokyo 105-8461 Japan. E-mail: aikou7000@gmail.com.

The authors declare no conflicts of interest.

Copyright © 2014 by IGCS and ESGO
ISSN: 1048-891X
DOI: 10.1097/IGC.0000000000000289

Copyright © 2014 by IGCS and ESGO. Unauthorized reproduction of this article is prohibited.

Ovarian cancer is made up of several different histological subtypes, and it is clear that these represent different diseases with distinct biology, pathogenesis, and clinical behavior. However, to date, they have all been treated in the same way. As understanding of the differences increases, this is no longer a rational approach. Most women included in clinical trials have high-grade serous (HGS) ovarian cancer, and it cannot be assumed that the results of these trials are applicable to women with other histotypes. Clear cell carcinoma of the ovary (CCC) is more likely to be detected at an early stage than HGS cancers, and when confined within the ovary, the prognosis is good. However advanced disease is associated with a very poor prognosis and resistance to standard treatment.[1] Histotype-specific trials and treatment protocols are required. Performing these trials and generating the evidence will require considerable international collaboration.

EPIDEMIOLOGY

The CCC is a histologic subtype of epithelial ovarian cancer with a distinct clinical behavior. There are marked geographic differences in the prevalence of CCC. In North America and Europe, CCC is the third most common histologic subtype of epithelial ovarian cancer, with an estimated prevalence of 1% to 12%.[2] Recent Surveillance, Epidemiology, and End Results data revealed that the incidences of CCC in women living in United States were 4.8% in whites, 3.1% in blacks, and 11.1% in Asians.[3] In Japan, the prevalence of CCC is higher than in western countries, although the reason for this remains unknown.[1] The annual report of the Japanese Gynecologic Cancer Committee showed an increasing incidence of CCC as a proportion of all epithelial ovarian cancers (Fig. 1),[4] now making up more than 25% of epithelial ovarian cancers in Japan.

The incidence of thromboembolic complications in CCC, such as deep venous thrombosis and pulmonary embolism, is reported to be higher than other epithelial ovarian cancers (16.9%–27.3% vs 0%–6.8%) and is considered as an independent prognostic factor.[5,6]

PATHOLOGY

Gross

Most CCCs are unilateral. Typically, the sectioned surface of the tumor reveals a unilocular cyst with 1 or more solid, yellow nodules protruding into the cyst. Cysts may contain watery, mucinous fluid or brownish "chocolate-colored" fluid. Multilocular cysts are less common, and occasional tumors are predominantly solid. The mean size of CCC is 15 cm.

Microscopic

The CCC is composed of glycogen-containing cells with abundant clear cytoplasm and hobnail cells. Many tumors also contain cells with granular eosinophilic cytoplasm. Nuclei are often eccentrically placed, with rounded-to-angulated contours. Hobnail cells have scant cytoplasm and enlarged, bulbous, hyperchromatic nuclei that protrude into tubule and cyst lumens. Bland and flattened cuboidal cells may line cysts or glands. It may arise within an endometriotic cyst, and benign endometriotic lesions or atypical endometriosis may be seen adjacent to CCC. Occasionally, CCCs also present in association with adenofibromatous, clear cell borderline tumors.

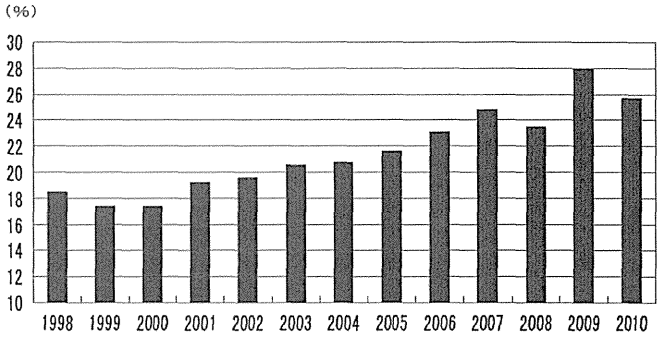


FIGURE 1. The rate of CCC among all epithelial ovarian cancers in Japan (annual reports on Japanese Gynecologic Cancer Committee).

The CCC exhibits high-grade nuclear features, although a spectrum of nuclear atypia may be present. The CCCs have traditionally been considered to be of high grade, but mitotic figures are relatively uncommon compared with other ovarian carcinomas.

Architectural patterns include tubulocystic, papillary, solid, and mixtures of them. Tubulocystic areas include tubules and cysts that are lined by flat-to-cuboidal cells with variable atypia and scattered hobnail cells (Fig. 2A). Papillary areas contain papillae that are small and round in comparison with those in serous carcinoma (Fig. 2B). The fibrovascular cores may be filled with either fibromatous, myxoid, spherulelike mucoid, or hyalinized basement membrane-type material. Solid areas are composed of sheets of polyhedral cells with clear cytoplasm (Fig. 2C).

Mixed subtypes of epithelial carcinomas are found. However, these should be considered as HGS tumors.[7]

Immunohistochemistry

The differential keratin profile is CK7+/CK20−, although CK7 may be focal in approximately 10% of cases.[8] In general, CCCs are negative for estrogen receptor, progesterone receptor, and *WT1*. Hepatocyte nuclear factor-1ß is a relatively new marker that is positive in CCC.[9] WT1 is useful in distinguishing CCC from mixed serous/clear cell tumors as it is typically positive in the latter.

Molecular Biology and Genetics

Unlike HGS tumors, CCCs are generally p53 wild type and have a lower frequency of *BRCA1* and *BRCA2* mutations.[10] The most frequent alterations are *ARID1A* and *PIK3CA* mutations.[11,12] *ARID1A* encodes the protein *BAF250a*, which is integral in the SWI-SNF chromatin remodeling complex. *ARID1A* mutations are seen in 40% to 60% of CCCs, but not in HGS carcinomas. In general, loss of *BAF250a* expression correlates with mutational status. *PIK3CA* mutations are seen in approximately 40% of clear cell tumors. Amplification and overexpression of the antiapoptotic protein, *PPM1D*, is seen in 10% of CCCs,[13] and mutation of *PPP2R1A* has been reported

© 2014 IGCS and ESGO

Copyright © 2014 by IGCS and ESGO. Unauthorized reproduction of this article is prohibited.

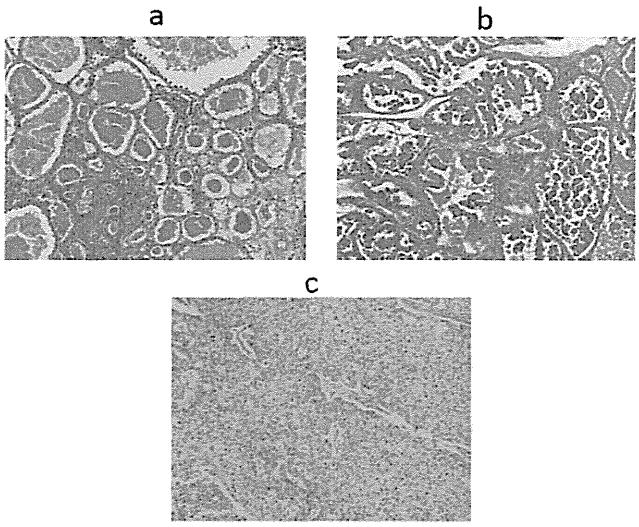


FIGURE 2. Microscopic findings of CCC. Tubulocystic areas include tubules and cysts that are lined by flat-to-cuboidal cells with variable atypia and scattered hobnail cells (A). Papillary areas contain papillae that are small and round in comparison with those in serous carcinoma (B). The fibrovascular cores may be filled with either fibromatous, myxoid, spherulelike mucoid, or hyalinized basement membrane-type material. Solid areas are composed of sheets of polyhedral cells with clear cytoplasm (C).

in 7%, whereas KRAS has been reported in 5%.[14] The CCCs are not a uniform group. Tan et al demonstrated groups with distinct patterns of copy number aberration in a comparative genomic hybridization analysis,[15] which seems to have prognostic significance.

Gene Expression Analysis

Yamaguchi et al[16] identified the gene signature that distinguishes CCC from other types of ovarian cancer using a microarray data set of ovarian cancers. The signature consisted of 437 genes and was designated as the CCC signature, which is specific for CCC. A categorical analysis demonstrated that genes belonging to 3 categories—stress response, sugar metabolism, and coagulation—are frequently involved in this signature.

INITIAL TREATMENT

Appropriate surgical treatment, followed by systemic chemotherapy, is recommended as an initial treatment for patients with CCC. The standard surgical treatment for patients with CCC is the same as for other epithelial ovarian cancers and includes hysterectomy, bilateral salpingo-oophorectomy, omentectomy, pelvic and para-aortic lymphadenectomy, and cytoreductive surgery. The recommended regimen of postoperative chemotherapy is paclitaxel (175 mg/m^2) combined with carboplatin (AUC 5–7.5), given every 3 weeks for 6 cycles.

Surgery

Lymphadenectomy is important to detect whether lymph nodes are involved in CCC because the presence of lymph node metastasis is an independent prognostic factor[17,18] and may guide the need for adjuvant therapy in early disease. Several authors have examined the therapeutic role of lymphadenectomy in therapy for this disease. In the Multicenter Italian Trials in Ovarian Cancer (MITO 9) retrospective study, disease-free survival in patients who underwent lymphadenectomy was longer than in other patients ($P = 0.0001$), in both early (I/II) ($P = 0.0258$) and advanced (III/IV) ($P = 0.037$) stages.[18] Lymphadenectomy also prolonged overall survival (OS) in patients with advanced stage ($P = 0.0039$). However, previous other reports have failed to show a therapeutic benefit from lymphadenectomy.[17,19] Further study will be required to identify the impact of lymphadenectomy on a patient's outcome from CCC.

Cytoreductive surgery should be performed for patients with stage II, III, or IV disease. Takano et al[20] reported no significant prognostic difference between the patients who underwent optimal cytoreduction (<1 cm) and those who had residual disease of greater than 1 cm. Complete surgery with no residual macroscopic disease was the only independent prognostic factor (median progression-free survival, 7 vs 5 vs 39 months, respectively). In a study by the Gynecologic Oncology Group, the markedly poor prognosis of CCC was observed even when patients have small-volume disease.[21] These findings suggest that a maximal effort should be made to remove all gross disease in patients with CCC.

Unilateral salpingo-oophorectomy preserving contralateral normal ovary and uterus should be considered for patients desiring to remain fertile. Several studies have examined outcomes of fertility-sparing surgery in patients with stage I CCC.[22,23] A total of 23 IA patients underwent fertility-sparing surgery, and all patients, excluding one (4%), were alive without recurrence. In contrast, 6 (25%) of the 24 patients at stage IC relapsed after surgery. Therefore, fertility-sparing surgery should only be offered for patients with stage IA disease.

Adjuvant Therapy

All patients with CCC have traditionally received postoperative systemic chemotherapy. However, observation may be considered for patients with surgical stage IA disease, because survival for these women is greater than 95%.[20,22,23]

It is generally accepted that CCC is resistant to conventional platinum-based chemotherapy compared with HGS ovarian cancer. The variation in reported response rates may reflect heterogeneity in patients included with some older studies including those with mixed tumors that would now be considered to be HGS tumors. Combination chemotherapy with paclitaxel plus platinum (TC) is thought to yield a higher response rate than conventional platinum-based chemotherapy (22%–56% vs 11%–27%) and improved survival in patients with advanced CCC, especially for those with optimal cytoreduction,[24,25] although the addition of a taxane was not an independent prognostic factor in the MITO 9 study. Nevertheless, responses remain much lower than with HGS, and there is an urgent need for more effective therapies.

In a randomized phase II study, the Japanese Gynecologic Oncology Group compared irinotecan hydrochloride plus cisplatin (CPT-P) with TC.[26] Both regimens were tolerated well, and progression-free survival between the 2 groups was

© 2014 IGCS and ESGO

Copyright © 2014 by IGCS and ESGO. Unauthorized reproduction of this article is prohibited.