A Competing Risks Model with Time Varying Covariates for Estimation of Breast Cancer Risks in \textit{BRCA1} Families

Yun-Hee Choi\textsuperscript{1,*}, Hae Jung\textsuperscript{1}, Fode Tounkara\textsuperscript{2}, Saundra Buys\textsuperscript{3}, Mary Daly\textsuperscript{4}, Esther M. John\textsuperscript{5}, John Hopper\textsuperscript{6}, Irene Andrilis\textsuperscript{2}, Mary Beth Terry\textsuperscript{7}, Laurent Briollais\textsuperscript{2,8}

\textsuperscript{1,*}Department of Epidemiology and Biostatistics, University of Western Ontario, London, Canada
\textsuperscript{2}Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, Canada
\textsuperscript{3}University of Utah Health Sciences Center, Salt Lake City, Utah, USA
\textsuperscript{4}Fox Chase Cancer Center, Philadelphia, Pennsylvania, USA
\textsuperscript{5}Stanford University School of Medicine, Stanford, California, USA
\textsuperscript{6}School of Population and Global Health, The University of Melbourne, Carlton, Victoria, Australia
\textsuperscript{7}Mailman School of Public Health, Columbia University, New York, USA
\textsuperscript{8}Division of Biostatistics, Dalla Lana School of Public Health, University of Toronto, Canada

\* Yun-Hee.Choi@schulich.uwo.ca

Abstract

Mammographic screening and prophylactic surgery can potentially reduce breast cancer risks among mutation carriers of \textit{BRCA} families. The evaluation of these interventions is usually complicated by the fact that their effects may change over time and by the presence of competing risks. We propose a competing risks model that accounts for time-varying interventions and provide cause-specific penetrance estimates for breast and ovarian cancers in \textit{BRCA1} families. A shared frailty model is specified to account for familial residual dependence with an ascertainment correction through affected probands, which
accounts for competing risks and time-varying covariates (TVCs). Via simulation studies we demonstrate the good performances of our proposed approach in terms of bias and precision of the estimators of model parameters and cause-specific penetrances over different levels of familial correlations. We apply our new approach to 498 BRCA1 mutation carrier families recruited through the Breast Cancer Family Registry and illustrate the importance of our approach accounted for both competing risks and TVCs when estimating cause-specific penetrance of breast cancer. Breast and ovarian cancers; BRCA; Competing risks; Time-varying covariate; Frailty model; Mammography; Prophylactic surgery; Penetrance.

1 Introduction

Between 10-15% of all breast cancers (BCs) are caused by a hereditary predisposition [Aloraifi and others, 2015]. Hereditary Breast and Ovarian Cancer syndrome (HBOC) is an autosomal dominant disease characterized by germline pathogenic mutations in the BRCA1 and BRCA2 genes for the majority of cases. It is the most common cause of hereditary forms of both breast and ovarian cancer (OC) [Petrucelli and others, 2010]. The overall prevalence of BRCA1/2 mutations is estimated to be from 1 in 400 to 1 in 800 with a higher prevalence in the Ashkenazi Jewish population (1 in 40). Estimates of penetrance (cancer risk) for BRCA1/2 mutations vary considerably (Petrucelli et al., 2010). Previous large meta-analyses reported mean cumulative BC risks at age 70 of 57% for BRCA1 and 49% for BRCA2 mutation carriers [Chen and Parmigiani, 2007; Kuchenbaecker and others, 2017]. The OC risks were 40% for BRCA1 and 18% for BRCA2 mutation carriers. Mutation carriers are also at an elevated risk of developing contralateral breast cancer (CBC) after a previous unilateral BC [Kuchenbaecker and others, 2017]. A recent meta-analysis estimated the 5-year CBC risk at 15% for BRCA1 mutation carriers and 9% for BRCA2 mutation carriers after a first BC [Molina-Montes and others, 2014]. Risk prediction models can be used to assess these risks in BRCA1/2 mutation positive families. These statistical models can help health practitioners to guide women who could benefit from genetic counselling and also in their clinical management, which currently comprise intensified surveillance for early BC detection using multimodal imaging techniques or prophylactic surgery such as bilateral mastectomy for the risk of BC and risk-reducing bilateral salpingo-oophorectomy (rrBSO) for the risk of OC.
BC and OC risk prediction models aim at estimating a woman’s absolute risk of developing cancer either for a fixed horizon or for a woman’s remaining lifetime. Several risk prediction models have been proposed for HBOC, the most popular being the Gail model (Gail et al., 1989), the Claus model (Claus and others, 1994) and its extension eClaus (Narod and others, 1995), BRCAPRO (Mazzola and others, 2014), BOADICEA (Lee and others, 2014) and the Tyrer-Cuzick (or IBIS) (Tyrer and others, 2004). The relative performance of four commonly used models for breast cancer risk (IBIS, BOADICEA, BRCAPRO and Gail/BCRAT) have been recently validated (Terry et al., 2019a). These models have a number of limitations. They cannot estimate accurately the risks of multiple cancer events (such as BC or OC) within the same families and usually do not account for death as a competing risk when estimating BC or OC age-dependent risks. Moreover, risk factors that have time-dependent effects on cancer risks are usually ignored or considered fixed over the whole study period (e.g., hormonal factors, mammographic screening (MS), surgical interventions). However, assessing how these factors modify a woman’s BC and OC risks and how their effects vary with a woman’s age (e.g., age at a surgical intervention or age at which screening is introduced) is critical to guide clinical decisions.

Competing risks models for clustered failure times data have already been proposed by Gorfine and Hsu (2011), which extended the competing risks model of Prentice and others (1978) to incorporate the frailty variables to cause-specific hazards models for all the causes. In a subsequent paper, Gorfine and others (2014) showed through a simulation study that naively treating competing risks as independent right censoring events resulted in non-calibrated predictions of cancer risks, with the expected number of events overestimated. Recently, we have also proposed a competing risks approach for clustered family data applicable to successive time-to-event outcomes (i.e. the first and second cancer event could each have a competing risk event) (Choi and others, 2017). However, to our knowledge, none of these approaches was developed to include time varying covariates (TVCs).

In clinical setting, assessing the effect of TVCs is important especially when the follow-up duration is long. For example, we can consider a binary variable for a certain treatment occurring at a later period of the follow-up duration. If we code
this variable as time invariant covariate (TIC), the duration of treatment exposure
becomes much longer than the actual exposure. We lose the information that the sub-
ject was actually absent of its effect for most part of the follow-up period. This type of
TVC is referred to as permanent exposure (PE) (Keown-Stoneman and others 2018)
as its effect stays constant permanently since the time of the treatment exposure.
Keown-Stoneman and others (2018) proposed the formulation of TVC effect, which
decays over time with the rate parameter referred to as exponential decay (ED). Cox
and Oakes (1984) have an additional parameter that measures the converged effect
of TVC, referred to as Cox and Oaks (CO) type.

In this paper, our goal is to extend our previous competing risks approach (Choi
and others 2017) to the situation where the cause-specific hazard function for the
main event of interest, BC, can depend on TVCs such as MS or rrBSO. Another
critical extension is to propose an ascertainment correction that specifically accounts
for the fact that the BRCA1 families have been recruited through a proband affected
by either BC or OC before her study entry, or through an unaffected proband. Finally,
in this framework, residual familial correlation not due to the BRCA1 mutation is
modelled through a shared frailty. With our proposed approach, we have BC, OC and
death from other causes as competing events in BRCA1 mutation families. We also
demonstrated a very relevant application of our model to a large series of BRCA1
families, in particular, with an assessment of rrBSO. The possibility that rrBSO
prevents future BC has been the subject of some debate. Terry and others (2019) did
not find an association after accounting for the time-varying nature of the covariate.
There may be some benefit in rrBSO, however women may elect for rrBSO close to
menopause limiting the impact. Here we consider the impact of the timing of rrBSO
in addition to MS through both simulations and applied analyses.

2 Model

2.1 Shared frailty competing risk model with time-varying covariates

Consider data arising from $n$ independent families, with family $f$, $f = 1, \ldots, n$, each
family consisting of $n_f$ members, $i = 1, \ldots, n_f$. For family member $i$ in family $f$, we
denote by $T_{fi}$ and $C_{fi}$ the time to the first event time and the right censoring time,
respectively, and by $\delta_{f_i} \in \{1, \ldots, J\}$ the type of the first observed event among $J$ competing events and $\delta_{f_i} = 0$ if right censored. The observed time is then defined as $T_{f_i} = \min(T^o_{f_i}, C_{f_i})$. We denote by $z_{f_j}$ the shared frailty specific to the event $j$ ($j = 1, \ldots, J$) within family $f$. To allow covariates to vary over time, let $x_{f_i}(t)$ be the vector of TVCs at time $t$ for individual $i$ in family $f$ and $X_{f_i}(t) = \{x_{f_i}(u); 0 \leq u < t\}$ represent the covariate history up to time $t$. Then the cause-specific hazard function for event $j$ for individual $i$ from family $f$ conditional on the covariate history $X_{f_i}(t)$ and cause-specific familial frailty $z_{f_j}$ follows a proportional hazards regression model

$$h_{f_{ij}}(t|X_{f_i}(t), z_{f_j}) = \lim_{dt \to 0} \frac{1}{dt}P(t \leq T_{f_i} < t + dt, \delta_{f_i} = j|T_{f_i} \geq t, X_{f_i}(t), z_{f_j})$$

$$= h_{0j}(t)z_{f_j}e^{\beta^T_j X_{f_i}(t)},$$

(1)

where $h_{0j}(t)$ is the baseline hazard function and $\beta_j$ is the vector of the covariate effects related to event $j$. Assuming that the future values of covariates up to any time $t > u$ are not affected by the occurrence of any event at time $u$, we can define the overall survival function across all competing events conditional on the covariate history and frailties as

$$S_{f_i}(t|X_{f_i}(t), z_f) = \exp\left\{-\sum_{j=1}^{J} H_{f_{ij}}(t|X_{f_i}(t), z_{f_j})\right\},$$

(2)

where $z_f = \{z_{f_1}, \ldots, z_{f_J}\}$ and $H_{f_{ij}}(t|X_{f_i}(t), z_{f_j}) = \int_0^t h_{0j}(u)z_{f_j}e^{\beta^T_j X_{f_i}(u)}du$ is the cause-specific cumulative hazard function at time $t$.

Suppose a time varying covariate $x_{f_i}(t) = 0$ at $t < t_x$ and $1$ at $t \geq t_x$, where $t_x$ is the time that change in value of covariate occurred. We can describe the effect of the TVC that changes over time, denoted by $m(\cdot)$, in three different structures: PE, ED, and CO as follows,

$$m(X_{f_i}(t)) = \begin{cases} 
0 & \text{if } t < t_x \text{ (PE, ED, CO)} \\
\beta & \text{if } t \geq t_x \text{ (PE)} \\
\beta \exp\{-\eta(t-t_x)\} & \text{if } t \geq t_x \text{ (ED)} \\
\beta \exp\{-\eta(t-t_x)\} + \eta_0 & \text{if } t \geq t_x \text{ (CO)}
\end{cases},$$

where for time $t \geq t_x$, the effect of TVC stays at $\beta$ for PE, whereas it starts to decrease exponentially with a rate of $e^{-\eta}$ to $0$ for ED or to $\eta_0$ for CO. The $j$th cause-specific hazard and cumulative hazard function with TVC can be written as

$$h_{f_{ij}}(t|X_{f_i}(t), z_{f_j}) = h_{0j}(t)z_{f_j}\exp\{m(X_{f_i}(t))\},$$
\[
H_{f_j}(t|X_{f_i}(t), z_{f_j}) = \int_0^t h_{0j}(u)z_{f_j}\exp\{m(X_{f_i}(u))\}du,
\]
where calculation details for cause-specific cumulative hazard for PE, ED and CO models are specified in Web Appendix A.

### 2.2 Likelihood construction

Let \( \theta = \{h_{0j}(\cdot), \beta_j, k_j, \eta_j, \eta_{0j}, j = 1, \ldots, J\} \) be the vector of parameters involved in the model, which consists of baseline parameters for specifying baseline hazard functions, regression coefficient vector \( \beta_j, \eta_j \) and \( \eta_{0j} \), related to TVC effects, and frailty parameter \( k_j \) for each competing event \( j = 1, \ldots, J \). Then, the likelihood of the data from \( n \) families can be constructed simply by the product of the likelihoods of all families:

\[
L(\theta) = \prod_{f=1}^{n} L_f(\theta).
\]

Under the shared frailty competing risk model framework, the likelihood for family \( f \) is obtained by integrating over the frailty distribution:

\[
L_f(\theta) = \prod_{i=1}^{n_f} \int_0^\infty \cdots \int_0^\infty \left\{ \prod_{j=1}^{J} h_{f_{ij}}(t_{f_i}|X_{f_i}(t_{f_i}), z_{f_j})^{I(\delta_{f_i} = j)} \right\} \times S_{f_i}(t_{f_i}|X_{f_i}(t_{f_i}), z_{f_j})g(z_{f_1})\ldots g(z_{f_J})dz_{f_1}\ldots dz_{f_J}.
\]

To compute the integrals, we utilize Laplace transform \( \phi(\cdot) \) of the frailty distribution \( g(z_{f_j}) \) and its \( d \)th derivative, \( \phi(\cdot)^{(d)} \), which have the following expressions

\[
\phi(s) = \int_0^\infty e^{-sz}g(z)dz
\]

\[
\phi(s)^{(d)} = (-1)^d \int_0^\infty z^d e^{-sz}g(z)dz.
\]

With \( z_{f_j} \sim \text{Gamma}(k_j, \frac{1}{k_j}) \), they have closed form expressions:

\[
\phi(s) = \left(1 + \frac{s}{k_j}\right)^{-k_j}
\]

\[
\phi(s)^{(d)} = (-1)^d \frac{(k_j + d - 1)!}{k_j!k_j^{d-1}} \left(1 + \frac{s}{k_j}\right)^{-k_j-d}.
\]
Then the likelihood for family $f$ can be obtained as

$$L_f(\theta) = \int_{0}^{\infty} \cdots \int_{0}^{\infty} \prod_{i=1}^{n_f} \prod_{j=1}^{J} \left\{ z_{fj} h_{ij}(t_{fi} | X_{fi}(t_{fi})) \right\}^{\delta_{fi} = j} \times e^{-\sum_{j=1}^{J} z_{fj} H_{ij}(t_{fi} | X_{fi}(t_{fi}))} g(z_{f1}) \cdots g(z_{fJ}) dz_{f1} \cdots dz_{fJ}$$

$$= \prod_{i=1}^{n_f} \prod_{j=1}^{J} h_{ij}(t_{fi} | X_{fi}(t_{fi}))^{\delta_{fi} = j} (-1)^{-d_{fj}} \phi^{(d_{fj})} \left\{ \sum_{i=1}^{n_f} H_{ij}(t_{fi} | X_{fi}(t_{fi})) \right\}$$

$$= \prod_{i=1}^{n_f} \prod_{j=1}^{J} h_{ij}(t_{fi} | X_{fi}(t_{fi}))^{\delta_{fi} = j} \frac{(k_j + d_{fj} - 1)!}{k_j! k_j^{d_{fj} - 1}} \left\{ 1 + \sum_{i=1}^{n_f} H_{ij}(t_{fi} | X_{fi}(t_{fi})) \right\}^{-k_j - d_{fj}}, \quad (3)$$

where $d_{fj} = \sum_{i=1}^{n_f} I(\delta_{fi} = j)$ is the number of family members affected by event $j$.

2.3 Ascertainment correction

We correct for the ascertainment bias by implementing prospective likelihood approach [Choi and others, 2008]. Families are ascertained via the probands (indexed as $p$) who have at least one of the competing events (BC, OC or death from other causes) before their age at examination ($a_{fp}$). The reason we also consider death as an ascertainment event is that in our real application, a small number of probands were unaffected at study entry but died during the follow-up period.

For each family $f$, we divide the $L_f(\theta)$ by the probability of the proband being ascertained by her age at examination, $A_f(\theta) = P(T_{fp} \leq a_{fp} | X_{fp}(a_{fp}))$, which can be derived by

$$A_f(\theta) = 1 - \int \cdots \int e^{-\sum_{j=1}^{J} z_{fp} H_{fpj}(a_{fp} | X_{fp}(a_{fp}))} \cdot g(z_{f1}) \cdots g(z_{fJ}) dz_{f1} \cdots dz_{fJ}$$

$$= 1 - \prod_{j=1}^{J} \left\{ 1 + \frac{H_{fpj}(a_{fp} | X_{fp}(a_{fp}))}{k_j} \right\}^{-k_j}. \quad (4)$$

In our real data application, we also consider unaffected probands. The ascertainment correction for them is given by

$$A_f(\theta) = \prod_{j=1}^{J} \left\{ 1 + \frac{H_{fpj}(a_{fp} | X_{fp}(a_{fp}))}{k_j} \right\}^{-k_j}.$$

Therefore, the ascertainment corrected likelihood for all the families is expressed as

$$L_C(\theta) = \prod_{f=1}^{n} \frac{L_f(\theta)}{A_f(\theta)}.$$
Combining results from (3) and (4), we specify

\[ L_C(\theta) = \prod_{f=1}^n \left( \prod_{i=1}^{n_f} \prod_{j=1}^J \left\{ h_{fij}(t_{fi}|X_{fi}(t_{fi})) \right\} \right) ^{I(\delta_{fi} = j) (k_j + d_{fj} - 1)!} \left( \frac{1}{k_j} \right)^{-k_j} \left( \prod_{j=1}^J \left\{ \sum_{i=1}^{n_f} H_{fij}(t_{fi}|X_{fi}(t_{fi})) \right\} \right) \]

and the corresponding log-likelihood to obtain maximum likelihood estimates of the parameters as

\[ \ell_C(\theta) = \sum_{f=1}^n \log L_f(\theta) - \sum_{f=1}^n \log A_f(\theta) \]

\[ = \sum_{f=1}^n \left\{ \sum_{i=1}^{n_f} \sum_{j=1}^J I(\delta_{fi} = j) \log h_{fij}(t_{fi}|X_{fi}(t_{fi})) \right\} \]

\[ + \sum_{f=1}^n \left[ \sum_{j=1}^J \log \left( (k_j + d_{fj} - 1)! \right) - \log(k_j) - (d_{fj} - 1) \log(k_j) \right] \]

\[ - (k_j + d_{fj}) \log \left( 1 + \sum_{i=1}^{n_f} H_{fij}(t_{fi}|X_{fi}(t_{fi})) \right) \]

\[ - \sum_{f=1}^n \log \left[ 1 - \prod_{j=1}^J \left\{ 1 + \frac{h_{fij}(t_{fi}|\delta_{fi} = j)}{k_j} \right\} ^{-k_j} \right]. (5) \]

### 2.4 Cause-specific penetrance function with time-varying covariates

Our main interest is to estimate the \( j \)th cause-specific cumulative incidence function \( F_j(\cdot) \), also called cause-specific penetrance. We first express the conditional cause-specific penetrance given the random frailty variable \( z_{fj} \) as

\[ F_{f_j}(t|X_{f_j}(t), z_{fj}) = P(T_{f_i} \leq t, \delta_{fi} = j|X_{f_i}(t), z_{f_j}) \]

\[ = \int_0^t h_{f_j}(u|X_{f_j}(u), z_{f_j}) \exp \left\{ - \sum_{j=1}^J H_{f_j}(u|X_{f_j}(u), z_{f_j}) \right\} du. \]
Assuming \( z_{f_j} \sim \text{Gamma}(k_j, \frac{1}{k_j}) \) and there are only two competing events, we derive the marginal cause-specific penetrance from event 1 (\( j = 1 \)) as follows:

\[
F_{f_1}(t|X_{f_1}(t)) = \int_0^\infty \int_0^\infty \int_0^t h_{f_1}(u|X_{f_1}(u), z_{f_1}) S_{f_1}(u|X_{f_1}(u), z_{f_1}, z_{f_2}) g(z_{f_1}) g(z_{f_2}) du dz_{f_1} dz_{f_2} = \int_0^t h_{f_1}(u)(-1)\phi^{(1)}\{H_{f_1}(u|X_{f_1}(u))\} \phi\{H_{f_1}(u|X_{f_1}(u))\} du \]

\[
= \int_0^t h_{f_1}(u)\left(1 + \frac{H_{f_1}(u|X_{f_1}(u))}{k_1}\right)^{-k_1-1} \left\{1 + \frac{H_{f_2}(u|X_{f_1}(u))}{k_2}\right\}^{-k_2} du,
\]

where calculation details for PE, ED and CO models are specified in Web Appendix B.

### 2.5 Variance Estimation

The variance-covariance matrix of \( \hat{\theta} \) is estimated using a robust sandwich variance estimator,

\[
V(\hat{\theta}) = I_o(\theta)^{-1} J(\theta) I_o(\theta)^{-1},
\]

where \( I_o(\theta) \) is the observed information matrix and \( J(\theta) \) is the expected information matrix. They can be obtained by

\[
I_o(\theta) = -\frac{\partial^2 \ell_C(\theta)}{\partial \theta^T \partial \theta},
\]

\[
J(\theta) = \sum_f U_f(\theta) U_f^T(\theta),
\]

\[
U_f(\theta) = \frac{\partial \log L_f(\theta)}{\partial \theta} - \frac{\partial \log A_f(\theta)}{\partial \theta}.
\]

The variance estimates \( \hat{V}(\hat{\theta}) \) are obtained by evaluating \( I_o(\theta) \) and \( J(\theta) \) at the maximum-likelihood estimate \( \hat{\theta} \).

The robust variance estimator for the cause-specific penetrance estimate, \( F_j(t|\hat{\theta}) \), is obtained using Delta method:

\[
V(F_j(t|\hat{\theta})) = D_{\theta}^T(t) V(\hat{\theta}) D_{\theta}(t),
\]

where \( D_{\theta}(t) \) is the vector of partial derivatives of \( F_j(t|\hat{\theta}) \) with respect to \( \theta \) evaluated at \( \hat{\theta} \). The variance estimates \( \hat{V}(F_j(t|\hat{\theta})) \) are obtained by using \( \hat{V}(\hat{\theta}) \).
3 Simulation study

3.1 Simulation Study Design

We conducted simulation studies to assess the finite-sample properties of our proposed shared frailty competing risks model. We considered $J = 2$ competing events with a TVC affecting a single event. Our simulated datasets mimic $BRCA1$ mutation positive families from the Breast Cancer Family Registry (BCFR) used in our application with respect to family structure and inclusion criteria. True parameter values were obtained after fitting our model to the real data. For each dataset, 500 families were generated under PE, ED and CO TVC models, each with low, medium and high familial dependence, which corresponds to $k_1 = 7$ ($\tau = 0.07$), $k_1 = 3.5$ ($\tau = 0.13$) and $k_1 = 1$ ($\tau = 0.33$), respectively, where $\tau$ represents a Kendall’s $\tau$. A value close to 1 indicates higher dependence among the family relatives’ failure times. The parameter $k_2$ was fixed at the estimated value obtained from the real data analysis. All combinations of parameters can be found in Table 1. The model included the mutation status as a TIC affecting both events and MS as a TVC for event 1. Detailed steps of data generation are presented in Web Appendix C. For each scenario, the model parameters and penetrance estimators are evaluated based on 500 simulations by comparing bias, empirical standard error (ESE), average standard error (ASE) and empirical coverage probability (ECP). Bias is defined as the difference between mean estimate, $\hat{\beta}$ and the true value of the parameter, $\beta$; ESE is obtained by the standard deviation of the estimates over all simulations, $\sqrt{\sum_{i=1}^{B}(\hat{\beta}_i - \bar{\hat{\beta}})^2/(B-1)}$, where $B = 500$ is the number of simulations and $\hat{\beta}_i$ is the parameter estimate from simulation $i$, $i = 1, \ldots, B$; ASE is obtained by $\sum_{i=1}^{B}SE(\hat{\beta}_i)/B$, the average of robust standard errors (SEs) from each simulation. Finally, ECP is the proportion of times 95% confidence interval (CI) defined as $\hat{\beta}_i \pm Z_{0.975}SE(\hat{\beta}_i)$ include true value $\beta$ for $i = 1, \ldots, B$.

In addition, we also investigated the robustness of the proposed model to misspecification of the TVC function in our simulations. Bias and efficiency of the misspecified TVC function are evaluated in comparison to the true TVC model. Simulations results based on $n = 500$ families are presented below while Web Tables 1 and 2 include simulation results for $n = 1000$ families.
3.2 Simulation Results

The simulation results for the model parameter estimates are summarized in Table 1. Biases of the parameter estimates related to the baseline hazard function \((\rho_1, \lambda_1, \rho_2, \lambda_2)\) and regression coefficients \((\beta_{1\text{screen}}, \beta_{1\text{gene}}, \beta_{2\text{gene}})\) are negligible across all the TVC models and the levels of familial dependences. ASEs and ESEs are very close to each other and ECPs are within acceptable range, i.e., between 0.93 and 0.97. The frailty parameter estimates are more biased especially for event 2 and their ECP is lower than the nominal level, 0.95 (ranged between 0.80 and 0.90). We also observed that ASEs tend to be larger than ESEs in the CO model. Coverage probability for \(k_1\) was better than for \(k_2\) and the bias decreases with the level of familial dependence.

Table 2 summarizes the simulation results related to the penetrance estimators. While frailty parameter estimators suffer from bias, penetrance estimators by age 70 for both event 1, \(F_1(70; X)\), and event 2, \(F_2(70; X)\), performed well. The bias was negligible (< 1%) and the ECPs were close to the 0.95 nominal level and within acceptable range (between 0.93 and 0.97) regardless of the level of familial dependence. ASEs and ESEs agree with each other in PE model but ASEs tend to be slightly higher than ESEs in the ED and CO models.

Additional simulations were conducted to evaluate the robustness of the proposed model to misspecification of the TVC function. We generated datasets under each TVC model assumption considering a medium familial dependence level \((k_1 = 3.5)\) and then fitting the wrong TVC models to them. Web Tables 3 and 4 summarize the simulation results for penetrance estimates under TVC misspecification. As expected, fitting ED and CO models on the dataset generated under a PE TVC leads to minimal biases. However, we note that screening coefficient \(\beta_{1\text{screen}}\) is largely biased under the CO model. Web Table 3 shows the screening effect \(\beta_{1\text{screen}}\) is underestimated while \(\eta_0\) is overestimated. The overall effect on penetrance is however unbiased since the bias on these two parameters is in opposite direction. Fitting a CO model on ED-generated data does not result in any bias. In other situations where a simpler TVC model is fitted to more complex true TVC models, substantial biases are observed for the screened individuals. Therefore, in practice, it is necessary to fit all three models and select the best model according to the lowest AIC values. In our simulations we note that the correct model is selected about 88% of the time with this selection criteria. In Web Tables 1 and 2, we present additional simulation results for parameter and penetrance estimators for a larger number of families \(n = 1000\).
brief, when \( n = 1000 \) the bias is substantially lower for all parameters, especially the frailty parameters, and their ECPs greatly improve (0.88 \sim 0.93 \) for \( k_2 \). Similarly, penetrance estimators are less biased, i.e. less than 0.1%.

## 4 Application to \textit{BRCA1} Families from BCFR

### 4.1 Data

Our analyses focus on \textit{BRCA1} positive families recruited through the BCFR. The BCFR \cite{John2004} was established in 1995 with six participating sites from the USA, Australia and Canada including Ontario Cancer Care. The BCFR enrolled most of the families from 1996 to 2000 while continuing to recruit additional families satisfying its criteria. Families were included whenever they segregate \textit{BRCA1} or \textit{BRCA2} mutations, exhibit multiple cases of breast or ovarian cancer, are Ashkenazi Jewish ancestry or from specific racial and ethnic groups. For the population-based families, each family includes the proband, i.e. the initial member of the family to be identified, as well as the first and the second degree relatives. The data have extensive information on the family members including the ages of the breast/ovarian cancer diagnosis, study entry, surgeries, and mammographic screening as well as mutation status in \textit{BRCA1/2} gene. We restricted our data analyses to the \textit{BRCA1} families in the BCFR, which were identified from 498 probands including a total of 2,650 individuals. The descriptive statistics of the data are summarized in Table 3.

### 4.2 Analyses

The first primary BC is our event of interest and the first primary OC and death from other causes than BC or OC are the competing events in our analyses. Age was considered as the time scale, i.e. age at diagnosis for women with cancer (BC or OC), and age at last follow-up or death for women without a first BC or OC. We considered two TVCs in our the analysis: age at MS and age at rrBSO. Prophylactic bilateral mastectomy was considered as a censoring variable for BC. For MS, we considered up to three possible screening events. We only accounted for screening and surgery histories before any events of interest (BC, OC, death or censored). When the age at rrBSO was less than one year from the age at BC onset, we considered that both events occurred at the same time and thus rrBSO did not affect BC (\( n = 12 \)). The proportion of individuals with OC as first cancer is much lower than that of BC (6.9%
The proportion of subjects who underwent rrBSO among the BC cohort is 3% and at least one MS is 21.9%.

4.3 Selection of the TVC models

For both the MS and rrBSO variables, we used the AIC to select the best TVC model and evaluated the three models, i.e. PE, ED and CO, for each of them. The best model corresponds to the CO model for both the MS and rrBSO variables with an AIC of 19080.76 (Web Table 5). In this model, rrBSO and the 3 MSs were highly significant \((p < 0.005)\) based on the likelihood ratio test when comparing a model with rrBSO vs. no rrBSO (the 3 MSs included) and a model with the 3 MSs vs. no MS (rrBSO included), respectively. The description of the following results are based on this model. The form of the hazard function corresponding the best model and that of other TVC models are displayed in Web Figure 1. All the TVC models show a monotonous increase of the hazard with age but the ED and CO models both depict a sharp peak of the hazard at the age at MS or rrBSO.

4.4 Relative Risk of the BC and OC

The log relative risks associated with the BRCA1 mutation status, MSs, and rrBSO are presented in Table 4. The log relative risk of mutation status on BC is \(\beta_{1\text{gene}} = 2.26\) (SE = 0.13), which indicates that being a mutation carrier increases the cause-specific hazard for BC by approximately 9.54 times compared to non-carriers after adjustment for the three MSs, rrBSO and the residual familial correlation. The log relative risk of rrBSO on BC is \(\beta_{1\text{rrBSO}} = -4.75\) (SE = 1.46), which decreases the cause-specific hazard for BC rapidly at the time of rrBSO but this effect dramatically increases to \(\eta_{0\text{rrBSO}} = -0.45\) (SE = 0.24) with time. The MSs drastically increase the BC-specific hazard with associated log relative risks for first, second and third MS of \(\beta_{1\text{screen1}} = 3.40\) (SE = 0.27), \(\beta_{1\text{screen2}} = 3.87\) (SE = 0.51), \(\beta_{1\text{screen3}} = 3.79\) (SE = 0.92), respectively. Their effects decrease with time to \(\eta_{0\text{screen1}} = 0.30\) (SE = 0.15), \(\eta_{0\text{screen2}} = -0.59\) (SE = 0.47), \(\eta_{0\text{screen3}} = -0.24\) (SE = 0.50) for the first, second, third MS, respectively.

There is a positive association between BRCA1 mutation and the cause-specific hazard of developing OC, where being a mutation carrier increases the cause-specific hazard by approximately 4.41 times compared to non-carriers. The estimates of frailty parameters \(k_1\) and \(k_2\) are 3.15 (95% CI between 2.11 and 4.70) and 0.70 (95%
CI between 0.33 and 1.52). They correspond to Kendall's $\tau$ around 0.14 for BC and 0.42 for OC.

4.5 Cause-specific penetrance estimation.

Cause-specific penetrance estimates were obtained from the competing-risks model. When no TVCs are present, i.e., for women without MS and rrBSO, the penetrance estimate by age 70 is 61.5% (95% CI = (56.7, 66.0)% for BC, 12.0% (95% CI = (9.6, 15.0)% for OC, and 11.7% (95% CI = (9.5, 14.6)% for death among gene carriers. The relative contribution of each cancer event and death to the overall cumulative penetrance by age among mutation carriers is presented in Web Figure 2. The BC-specific penetrance estimates with respect to MS are displayed in Figure 1. The cumulative penetrance of BC by age 70 is 69.8% (95% CI = (62.5, 75.8)%), 59.9% (95% CI = (43.4, 77.4)% and 57.3% (95% CI = (39.7, 79.3)% for a woman having 1, 2 and 3 MSs, where the MS occurred at 35, 40 and 45 years, respectively. The BC-specific penetrance estimate with respect to rrBSO is displayed in Figure 2, where the left plot is derived from the competing-risks model and the right plot from a non-competing-risks model (i.e. BC as single event). Under the competing-risks model, for a woman having rrBSO at 40 years or 50 years, the penetrance of BC by age 70 is 50.2% (95% CI = (40.0, 60.3)% and 53.4% (95% CI = (46.2, 60.7)%), respectively. Under a non-competing-risks model, these estimates are 59.0% (95% CI = (45.5, 72.3)%) and 60.7% (95% CI = (50.6, 71.1)%), respectively.

5 Discussion

Members of *BRCA* mutation positive families are exposed to a very high risk of developing BC or OC as first cancer and the risk of BC is likely to depend on time-varying covariates such as MS and rrBSO in a complex manner. Most risk prediction models developed for these families do not account for competing risks nor for time-varying effects on BC. In this paper, we developed a flexible approach based on frailty models for modelling competing risks (BC, OC or death) in family data, where the risk of the first competing event (BC) could depend on time-varying covariates. Our model provides cause-specific cumulative incidence function that estimates age-specific risks of BC and OC with respect to fixed and time-varying covariates, accounting for
death as a competing event and residual familial correlation not due to the mutation segregating within the family.

Our simulation studies demonstrate the good performances of our approach in terms of bias and precision of the estimators of model parameters and cause-specific penetrances over different levels of familial correlations. The frailty-related parameter estimators had larger biases than other parameter estimators but these biases did not result in any biases of the cause-specific penetrances. This is a very important result since the cause-specific penetrance is used by genetic counsellors to guide clinical decisions such as prophylactic surgery or intensive screening for known mutation carriers or the decision to have genetic testing for unknown mutation carriers in BRCA families. Another important result is that, applying models with the wrong TVC function could also result in substantial biases of the parameter estimators when fitting a simpler model to a more complex time-varying function. It is therefore critical to select the correct TVC function to obtain accurate cause-specific penetrance estimates.

Our application to 498 BRCA1 mutation positive families from the BCFR illustrates the importance of accounting for both competing risks and TVCs when estimating cause-specific penetrance of BC among mutation carriers. This penetrance is clearly modified by interventions such as rrBSO and multiple MSs over a woman’s lifetime and our new model, which integrates these interventions, might help better evaluating their long-term and age-specific effects in BRCA families. For instance, we found that the cumulative BC risk by age 70 was 61.5% for a mutation carrier woman who did not have rrBSO or any MS. However, for a woman having rrBSO at 40 years or 50 years and no MS, the penetrance of BC by age 70 drops to 50.2% and 53.4%, respectively. In our data, there were 14 (0.5%) and 63 (2.4%) of women who had rrBSO before 40 or 50 years, respectively. Interestingly, under a non-competing risks model, these estimates are 59.0% and 60.7%, respectively, and lead to the conclusion that there is no preventive effect of rrBSO on BC. These results with a non-competing risks model assumption agree with those of Terry and others (2019), which also found no effect of rrBSO (HR=1.04, 95% CI=0.87-1.24). We also found little cumulative risk difference in women with rrBSO if they were non-mutation carriers similar to Terry and others (2019). In addition, with Terry and others (2019), we did find a decreased risk after rrBSO when using a competing risks model (Figure 2); the differences were stronger when the earlier rrBSO had occurred. Thus, the main difference between these two analyses were driven by the assumption of competing risks, which
we used here, and not assuming a permanent exposure (PE). Our results were similar
to Terry el al. (Web Appendix 6). Other differences between the two studies include
the sample population as we included retrospective BC and OC cases and did not
include data from the kConFab consortium (Terry and others 2016).

Our model assumes the TVCs as exogenous variables, i.e, the hazard function at
a specific time $t$ is influenced by the observed covariate history up to time $t$ in
the regression model, but the occurrence of BC in $[t, t + dt]$ is independent of the future
path of the covariate (Cortese and Andersen 2009). This assumption is realistic for
prophylactic rrBSO and scheduled MS in our application since the observation of
rrBSO and MS does not carry information about the status of BC; however, if the
MS was performed in symptomatic women, the MS may not be completely exogenous
since it could carry information about the status of BC. Even in that latter situation,
our inference is based on the likelihood conditional on the covariate process up to
the time $t$, so the future path of the covariate would not influence the occurrence of
BC. In the situation where the full path of the TVC is of research interest, e.g. even
after the event of interest, some statistical approaches, such as the joint modeling
of the TVCs as recurrent events and the cancer outcome as a terminal event, could
be proposed. We have recently developed such approach for family data however it
will require further extensions to be applicable to competing risks events (Choi and
others 2019).

Our model could also help evaluating more intervention options on BC risk, such
as combinations of rrBSO and MSs as well as the ages they could be introduced.
It could be further extended to account for additional competing risks events, e.g.
prophylactic mastectomy, and also to estimate the risks of successive cancer events
after a first BC or OC, for example following our previous work (Choi and others 2017).
Finally, we are planning to incorporate information on polygenic risk score
from known genetic variants (Kuchenbaecker and others 2017), that could modify
BC and OC risks by incorporating a kinship matrix into the cause-specific model for
BC and/or OC (Lakhal-Chaieb and others 2018). These future developments should
lead to a more comprehensive risk prediction model applicable to BRCA families as
well as other families with increased genetic risks.
Acknowledgements

This work was supported by grant UM1 CA164920 from the USA National Cancer Institute. The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the Breast Cancer Family Registry (BCFR), nor does mention of trade names, commercial products, or organizations imply endorsement by the USA Government or the BCFR. This research was also supported by two grants from the Canadian Institutes of Health Research (MOP 126186 & 110053), an Interdisciplinary Health Research Team award from the Canadian Institutes of Health Research (Grant # 43821), a grant from the Canadian Breast Cancer Foundation (BC-RG-15-2 competition), and Discovery Grants from the Natural Sciences and Engineering Research Council of Canada.

Conflict of Interest: None declared.

References

Aloraifi, F, Boland, M R, Green, A J and Geraghty, J G. (2015). Gene analysis techniques and susceptibility gene discovery in non-brca1/brca2 familial breast cancer. Surgical Oncology 24, 100–109.

American College of Obstetricians and Gynecologists. (2017). Practice bulletin no 182: Hereditary breast and ovarian cancer syndrome. Obstetrics and Gynecology 130, e110–e126. doi: 10.1097/aog.0000000000002296.

Chen, S and Parmigiani, G. (2007). Meta-analysis of brca1 and brca2 penetrance. Journal of clinical oncology 25, 1329–1333.

Choi, Y-H, Briollais, L, Win, A K, Hopper, J, Buchanan, D, Jenkins, M and Lakhal-Chaieb, L. (2017). Modelling of successive cancer risks in lynch syndrome families in the presence of competing risks using copulas. Biometrics 73, 271–282.

Choi, Y-H, Jacqmin-Gadda, H, Krol, A, Parfrey, P, Briollais, L and Rondeau, V. (2019). Joint nested frailty models for clustered recurrent and terminal events: An application to colonoscopy screening visits and colorectal cancer risks in lynch syndrome families. Statistical Methods in Medical Research, In press.
Choi, Y-H, Kopciuk, K and Briollais, L. (2008). Estimating disease risks associated with mutated genes in family-based designs. *Human Hereditary* **66**, 238–251.

Claus, E B, Risch, N and Thompson, W D. (1994). Autosomal dominant inheritance of early-onset breast cancer. Implications for risk prediction. *Cancer* **73**, 643–651.

Cortese, G and Andersen, P K. (2009). Competing risks and time-dependent covariates. *Biometrical Journal* **51**, 138–158.

Gorfine, M and Hsu, L. (2011). Frailty-based competing risks model for multivariate survival data. *Biometrics* **67**, 415–426.

Gorfine, M, Hsu, L, Zucker, DM and Parmigiani, G. (2014). Calibrated predictions for multivariate competing risks models. *Lifetime Data Analysis* **20**, 234–51.

John, E M, J, Hopper, Beck, J, Knight, J, Neuhausen, S, Senie, R and others. (2004). The breast cancer family registry: an infrastructure for cooperative multinational, interdisciplinary and translational studies of the genetic epidemiology of breast cancer. *Breast Cancer Research* **6**, R375–R389.

Keown-Stoneman, C, Horrocks, J and Darlington, G. (2018). Exponential decay for binary time-varying covariates in cox models. *Statistics in Medicine* **37**, 776–788.

Kuchenbaecker, K B, Hopper, J L, Barnes, D R, Philips, K A, Moolij, T M, Roos-Blom, M J and others. (2017a). Risks of breast, ovarian, and contralateral breast cancer for *brca1* and *brca2* mutation carriers. *JAMA* **317**, 2402–2416.

Kuchenbaecker, K B, McGuffog, L, Barrowdale, D, Lee, A, Soucy, P, Dennis, J and others. (2017b). Evaluation of polygenic risk scores for breast and ovarian cancer risk prediction in *brca1* and *brca2* mutation carriers. *Journal of the National Cancer Institute* **109**(7), djw302.

Lakhal-Chaieb, L, Simard, J and Bull, S. (2018). Sequence kernel association test for survival outcomes in the presence of a non-susceptible fraction. *Biostatistics*, In press.
Lee, A J, Cunningham, A P, Kuchenbaecker, K B, Mavaddat, N, Easton, D F, Antoniou, A C and others. (2014). Boadicea breast cancer risk prediction model: updates to cancer incidences, tumour pathology and web interface. British Journal of Cancer 110, 535–545.

Mazzola, E, Chipman, J, Cheng, S C and Parmigiani, G. (2014). Recent br-capro upgrades significantly improve calibration. Cancer Epidemiology, Biomarkers & Prevention 23, 1689–1695.

Molina-Montes, E, Perez-Nevot, B, Pollan, M, Sanchez-Cantalejo, E, Espin, J and Sanchez, M J. (2014). Cumulative risk of second primary contralateral breast cancer in brca1/brca2 mutation carriers with a first breast cancer: a systematic review and meta-analysis. The Breast 23, 721–742.

Narod, S A, Goldgar, D, Cannon-Albright, L, Weber, B, Moslehi, R, Ives, E, Lenoir, G and Lynch, H. (1995). Risk modifiers in carriers of brca1 mutations. International Journal of Cancer 64, 394–398.

Petrucelli, N, Daly, M B and Feldman, G L. (2010). Hereditary breast and ovarian cancer due to mutations in brca1 and brca2. Genetics in Medicine 12, 245–249.

Prentice, R L, Kalbfleisch, J D, Peterson, A V, Flournoy, N, T, Farewell V, and Breslow, N. (1978). The analysis of failure times in the presence of competing risks. Biometrics 34, 541–554.

Terry, M B, Daly, M B, Phillips, K A, Ma, X, Zeinomar, N, Leoce, N and others. (2019). Risk-reducing oophorectomy and breast cancer risk across the spectrum of familial risk. Journal of the National Cancer Institute 111, 331–334.

Terry, M B, Phillips, K-A, Daly, M B, John, E M, Andrilis, I L, Buys, S S and others. (2016). Cohort profile: the breast cancer prospective family study cohort (prof-sc). International Journal of Epidemiology 45, 683–692.

Tyrer, J, Duffy, S W and Cuzick, J. (2004). A breast cancer prediction model incorporating familial and personal risk factors. Statistics in Medicine 23, 1111–1130.
Table 1: Empirical parameter estimates from the competing risks model with a time varying covariate (TVC) under low ($k_1 = 7$), medium ($k_1 = 3.5$) and high ($k_1 = 1$) familial dependence; permanent exposure (PE), exponential decay (ED) or Cox and Oaks (CO) models are considered for TVC. For each scenario, the mean bias, empirical standard error (ESE), average standard error (ASE) and estimated 95% coverage probability (ECP) are obtained from 500 replicates each with $n = 500$ families.

| TVC model | True $k_1 = 7, \tau = 0.07$ | True $k_1 = 3.5, \tau = 0.13$ | True $k_1 = 1, \tau = 0.33$ |
|-----------|-------------------------------|-------------------------------|-------------------------------|
|           | value | Bias | ESE | ASE | ECP | value | Bias | ESE | ASE | ECP | value | Bias | ESE | ASE | ECP |
| PE        | log($\lambda_1$) | -4.83 | -0.01 | 0.06 | 0.95 | -4.83 | 0.00 | 0.06 | 0.95 | -4.83 | 0.00 | 0.06 | 0.95 |
|           | log($\mu_1$) | 0.88 | 0.00 | 0.03 | 0.94 | 0.88 | 0.00 | 0.03 | 0.93 | 0.88 | 0.00 | 0.03 | 0.96 |
|           | log($\lambda_2$) | -4.96 | -0.01 | 0.09 | 0.95 | -4.96 | -0.02 | 0.10 | 0.94 | -4.96 | -0.01 | 0.09 | 0.96 |
|           | log($\mu_2$) | 1.12 | 0.00 | 0.07 | 0.95 | 1.12 | 0.00 | 0.07 | 0.95 | 1.12 | 0.00 | 0.06 | 0.96 |
|           | $\beta_{screen}$ | 0.67 | 0.01 | 0.11 | 0.95 | 0.67 | 0.00 | 0.10 | 0.96 | 0.67 | 0.00 | 0.11 | 0.96 |
|           | $\beta_{gene}$ | 1.95 | 0.01 | 0.12 | 0.95 | 1.95 | 0.01 | 0.12 | 0.96 | 1.95 | 0.01 | 0.12 | 0.94 |
|           | log($k_1$) | 1.95 | 0.24 | 1.08 | 0.99 | 1.95 | 0.23 | 1.08 | 0.99 | 1.95 | 0.22 | 1.06 | 0.99 |
|           | log($k_2$) | 1.06 | 0.62 | 2.17 | 0.80 | 1.06 | 0.72 | 2.20 | 0.84 | 1.06 | 0.61 | 2.05 | 0.86 |
| ED        | log($\lambda_1$) | -4.83 | -0.01 | 0.05 | 0.96 | -4.83 | 0.00 | 0.06 | 0.95 | -4.83 | 0.00 | 0.06 | 0.96 |
|           | log($\mu_1$) | 0.83 | 0.00 | 0.03 | 0.96 | 0.83 | 0.00 | 0.03 | 0.96 | 0.83 | 0.00 | 0.03 | 0.96 |
|           | log($\lambda_2$) | -4.96 | 0.00 | 0.09 | 0.95 | -4.96 | -0.01 | 0.09 | 0.96 | -4.96 | -0.01 | 0.09 | 0.95 |
|           | log($\mu_2$) | 1.08 | 0.00 | 0.06 | 0.95 | 1.08 | 0.00 | 0.06 | 0.95 | 1.08 | 0.00 | 0.06 | 0.95 |
|           | $\beta_{screen}$ | 1.87 | 0.03 | 0.25 | 0.94 | 1.87 | 0.01 | 0.25 | 0.95 | 1.87 | 0.03 | 0.24 | 0.94 |
|           | $\beta_{gene}$ | 1.86 | 0.00 | 0.12 | 0.96 | 1.86 | 0.01 | 0.11 | 0.95 | 1.86 | 0.01 | 0.11 | 0.94 |
|           | log($k_1$) | 1.95 | 0.23 | 0.99 | 0.93 | 1.95 | 0.28 | 0.94 | 0.97 | 2.00 | 0.23 | 0.24 | 0.96 |
|           | log($k_2$) | 1.18 | 0.51 | 2.04 | 0.85 | 1.18 | 0.53 | 1.70 | 0.84 | 1.18 | 0.48 | 1.47 | 0.84 |
|           | log($\eta$) | -1.28 | 0.02 | 0.32 | 0.94 | -1.28 | 0.00 | 0.33 | 0.94 | -1.28 | 0.03 | 0.30 | 0.94 |
| CO        | log($\lambda_1$) | -4.83 | 0.00 | 0.05 | 0.95 | -4.83 | 0.00 | 0.05 | 0.94 | -4.83 | 0.00 | 0.05 | 0.96 |
|           | log($\mu_1$) | 0.83 | 0.00 | 0.03 | 0.94 | 0.83 | 0.00 | 0.03 | 0.96 | 0.83 | 0.00 | 0.03 | 0.97 |
|           | log($\lambda_2$) | -4.96 | 0.00 | 0.07 | 0.95 | -4.96 | 0.00 | 0.07 | 0.97 | -4.96 | 0.00 | 0.08 | 0.95 |
|           | log($\mu_2$) | 1.07 | 0.00 | 0.05 | 0.96 | 1.07 | 0.00 | 0.05 | 0.97 | 1.07 | 0.00 | 0.05 | 0.96 |
|           | $\beta_{screen}$ | 1.52 | 0.04 | 0.32 | 0.96 | 1.52 | 0.04 | 0.33 | 0.94 | 1.52 | 0.02 | 0.32 | 0.96 |
|           | $\beta_{gene}$ | 2.08 | 0.01 | 0.10 | 0.94 | 2.08 | 0.01 | 0.10 | 0.95 | 2.08 | 0.01 | 0.09 | 0.96 |
|           | log($k_1$) | 1.95 | 0.20 | 0.74 | 0.91 | 1.25 | 0.10 | 0.39 | 0.46 | 0.91 | 0.00 | 0.18 | 0.22 |
|           | log($k_2$) | 1.26 | 0.38 | 1.15 | 0.86 | 1.26 | 0.35 | 1.40 | 0.90 | 1.26 | 0.36 | 1.10 | 0.87 |
|           | log($\eta$) | -0.18 | -0.02 | 0.50 | 0.90 | -0.18 | 0.01 | 0.50 | 0.91 | -0.18 | -0.03 | 0.48 | 0.91 |
|           | $\eta_0$ | 0.21 | -0.02 | 0.12 | 0.95 | 0.21 | -0.01 | 0.12 | 0.96 | 0.21 | -0.02 | 0.12 | 0.95 |
Table 2: Empirical penetrance estimates by age 70 for the competing risks model with a time varying covariate (TVC) under low ($k_1 = 7$), medium ($k_1 = 3.5$) and high ($k_1 = 1$) familial dependence; permanent exposure (PE), exponential decay (ED) or Cox and Oaks (CO) models are considered for TVC; $F_1(70; S, G)$ and $F_2(70; S, G)$ are cause-specific penetrance estimators (%) by age 70 for event 1 and event 2, respectively, given screening (S) and mutation status (G), and screening time at age 35 if $S = 1$. For each scenario, the mean bias, empirical standard error (ESE), average standard error (ASE) and estimated 95% coverage probability (ECP) are obtained from 500 replicates each with $n = 500$ families.

| TVC model | $k_1 = 7, \tau = 0.07$ | $k_1 = 3.5, \tau = 0.13$ | $k_1 = 1, \tau = 0.33$ |
|-----------|-----------------|-----------------|-----------------|
|           | True value Bias | True value Bias | True value Bias |
| PE        |                |                |                |
| $F_1(70; S = 0, G = 0)$ | 12.56 -0.10 1.38 1.36 0.95 | 12.45 0.01 1.33 1.40 0.94 | 11.93 0.07 1.48 1.45 0.94 |
| $F_1(70; S = 1, G = 0)$ | 21.92 -0.01 2.45 2.45 0.94 | 21.58 0.02 2.37 2.48 0.95 | 20.99 0.13 2.49 2.50 0.96 |
| $F_1(70; S = 0, G = 1)$ | 56.52 -0.33 3.20 3.18 0.94 | 54.51 0.12 3.39 3.42 0.94 | 46.80 -0.02 3.84 3.92 0.95 |
| $F_1(70; S = 1, G = 1)$ | 75.63 -0.23 3.75 3.74 0.94 | 72.59 0.03 4.08 4.06 0.94 | 61.08 -0.04 4.61 4.79 0.94 |
| ED        |                |                |                |
| $F_1(70; S = 0, G = 0)$ | 4.73 -0.08 0.82 0.85 0.94 | 4.73 -0.08 0.87 0.85 0.93 | 4.74 -0.05 0.79 0.88 0.95 |
| $F_1(70; S = 1, G = 0)$ | 4.45 -0.08 0.77 0.80 0.94 | 4.45 -0.08 0.82 0.80 0.93 | 4.49 -0.05 0.75 0.83 0.95 |
| $F_1(70; S = 0, G = 1)$ | 9.68 0.04 1.16 1.15 0.94 | 9.85 -0.04 1.16 1.18 0.95 | 10.52 0.02 1.29 1.28 0.95 |
| $F_1(70; S = 1, G = 1)$ | 7.12 0.01 0.91 0.89 0.94 | 7.42 -0.04 0.91 0.92 0.95 | 8.56 0.00 1.04 1.04 0.95 |
| CO        |                |                |                |
| $F_1(70; S = 0, G = 0)$ | 13.55 -0.05 1.39 1.42 0.94 | 13.42 -0.02 1.41 1.44 0.94 | 12.82 -0.04 1.47 1.47 0.94 |
| $F_1(70; S = 1, G = 0)$ | 15.49 0.03 1.64 1.64 0.94 | 15.32 0.05 1.62 1.66 0.94 | 14.54 0.00 1.61 1.68 0.97 |
| $F_1(70; S = 0, G = 1)$ | 55.65 -0.28 2.70 3.07 0.97 | 53.68 -0.05 3.03 3.27 0.96 | 46.14 0.12 3.56 3.68 0.96 |
| $F_1(70; S = 1, G = 1)$ | 60.49 -0.10 2.99 3.33 0.97 | 58.24 0.10 3.26 3.54 0.97 | 49.69 0.21 3.67 3.94 0.96 |
|           |                |                |                |
| $F_2(70; S = 0, G = 0)$ | 5.39 0.01 0.90 0.91 0.95 | 5.39 -0.07 0.86 0.92 0.95 | 5.41 -0.05 0.85 0.93 0.96 |
| $F_2(70; S = 1, G = 0)$ | 5.26 0.01 0.87 0.89 0.95 | 5.26 -0.07 0.83 0.89 0.95 | 5.28 -0.06 0.83 0.91 0.95 |
| $F_2(70; S = 0, G = 1)$ | 11.38 0.05 1.18 1.22 0.96 | 11.57 0.04 1.29 1.24 0.95 | 12.34 -0.05 1.34 1.35 0.95 |
| $F_2(70; S = 1, G = 1)$ | 9.97 0.01 1.05 1.09 0.96 | 10.22 0.01 1.12 1.12 0.95 | 11.20 -0.07 1.20 1.22 0.95 |
Table 3: Descriptive statistics for *BRCA1* positive families.

|                      | Breast Cancer | Ovarian Cancer | Death | Unaffected | Total |
|----------------------|---------------|----------------|-------|------------|-------|
| **N(%)**             | 924 (34.9%)   | 182 (6.9%)     | 958 (36.2%) | 586 (22.1%) | 2650  |
| **N(%) of probands** | 391 (78.5%)   | 43 (8.6%)      | 5 (1.0%)  | 59 (11.9%)  | 498   |
|                      | 386 (77.5%)   | 31 (6.2%)      | 0 (0%)    | 81 (16.3%)  | 498   |

**Event age**

- **mean (SD)**: 44.2 (12.0), 53.0 (11.5), 70.5 (17.9), 50.9 (16.2), 55.8 (19.1)
- **min, max**: 21.0, 86.0, 28.0, 89.0, 18.5, 102.5, 18.1, 95.0, 18.1, 102.5

**BRCA mutation**

- Noncarrier: 29 (3.1%), 4 (2.2%), 14 (1.5%), 229 (39.1%), 276 (10.4%)
- Carrier: 483 (52.3%), 55 (30.2%), 16 (1.7%), 192 (32.8%), 746 (28.2%)
- Untested: 412 (44.6%), 123 (67.6%), 928 (96.9%), 165 (28.2%), 1628 (61.4%)

**# of mammographic screening**

- 0: 722 (78.1%), 158 (86.8%), 944 (98.5%), 257 (43.9%), 2081 (78.5%)
- 1: 160 (17.3%), 19 (10.4%), 7 (0.7%), 174 (29.7%), 360 (13.6%)
- 2: 31 (3.4%), 4 (2.2%), 3 (0.3%), 63 (10.8%), 101 (3.8%)
- 3+: 11 (1.2%), 1 (0.5%), 4 (0.4%), 92 (15.7%), 108 (4.1%)

**rrBSO**

- 28 (3.0%), 0 (0%), 9 (0.9%), 129 (22.0%), 166 (6.3%)

rrBSO stands for risk-reducing bilateral salpingo-oophorectomy.
SD stands for standard deviation.
Table 4: Parameter estimates associated with BC in the *BRCA1* families from BCFR based on the model with or without competing risks (OC and death) assuming CO model for mammography screening and CO for risk reducing bilateral salpingooophorectomy.

|                         | Competing risks model |                     | No competing risks model |                     |
|-------------------------|-----------------------|---------------------|--------------------------|---------------------|
|                         | Estimate   | SE     | *p*-value | Estimate   | SE     | *p*-value |
| $\beta_{1\text{gene}}$ | 2.255      | 0.128  | 0.000     | 2.218      | 0.132  | 0.000     |
| $\beta_{1\text{screen}}$ | 3.399      | 0.267  | 0.000     | 3.374      | 0.266  | 0.000     |
| $\beta_{1\text{screen}}$ | 3.870      | 0.510  | 0.000     | 3.843      | 0.510  | 0.000     |
| $\beta_{1\text{screen}}$ | 3.790      | 0.922  | 0.000     | 3.956      | 0.899  | 0.000     |
| $\beta_{1\text{rrBSO}}$ | -4.749     | 1.464  | 0.001     | -4.029     | 1.102  | 0.000     |
| log($\eta_{\text{screen}}$) | 1.402      | 0.253  | 0.000     | 1.481      | 0.249  | 0.000     |
| log($\eta_{\text{screen}}$) | 0.614      | 0.379  | 0.105     | 0.670      | 0.388  | 0.085     |
| log($\eta_{\text{screen}}$) | 1.632      | 1.002  | 0.103     | 1.675      | 1.005  | 0.096     |
| $\eta_{0\text{screen}}$ | 0.303      | 0.145  | 0.037     | 0.290      | 0.154  | 0.060     |
| $\eta_{0\text{screen}}$ | -0.594     | 0.466  | 0.202     | -0.602     | 0.470  | 0.200     |
| $\eta_{0\text{screen}}$ | -0.240     | 0.497  | 0.629     | -0.364     | 0.529  | 0.492     |
| log($\eta_{\text{rrBSO}}$) | 0.131      | 0.680  | 0.848     | 0.194      | 0.814  | 0.812     |
| $\eta_{0\text{rrBSO}}$ | -0.451     | 0.238  | 0.058     | -0.334     | 0.268  | 0.212     |
| log($k_1$)               | 1.148      | 0.204  | 0.000     | 1.074      | 0.216  | 0.000     |
| log($k_2$)               | 0.866      | 0.393  | 0.027     |             |        |           |
| $\beta_{2\text{gene}}$  | 1.483      | 0.229  | 0.000     |             |        |           |
| $\beta_{3\text{gene}}$  | -0.351     | 0.140  | 0.012     |             |        |           |
| -loglik                  | 9517.38    |        | 3862.46   |             |        |           |
| -loglik*                 | 9525.00    |        | 3866.32   |             |        |           |
| *p*-value                | 0.002      |        | 0.052     |             |        |           |

† based on the null model without rrBSO
* testing for rrBSO effect comparing to the null model based on the likelihood ratio test with df = 3
Figure 1: Breast cancer-specific penetrance estimates for mutation carriers with respect to multiple mammographic screenings (MSs). The red line represents a screened woman and the blue line a woman not screened. The dashed lines represent the 95% confidence intervals. The plot on the left displays the penetrance of a woman screened at 35 years vs. non screened. The centre plot displays the penetrance of a woman screened at 35 and 40 years vs. non screened. The plot on the right displays the penetrance of a woman screened at 35, 40 and 45 years vs. non screened.
Figure 2: Breast cancer-specific penetrance estimates for mutation carriers with respect to rrBSO. The left plot is derived from the competing-risks model and the right plot from a non-competing-risks model (i.e. BC as a single event). The black line represents a woman who did not have rrBSO, the green line a woman who had rrBSO at age 40 years and the blue line a woman who had rrBSO at age 50 years. The dashed lines represent the 95% confidence intervals.
Supplementary Materials

Web Appendix A:
Derivation of cumulative hazard function with a time-varying covariate

For the time varying covariate \( x_f_i(t) = 0 \) at \( t < t_x \) and \( 1 \) at \( t \geq t_x \), where \( t_x \) is the time that change in value of time varying covariate occurred. The \( j \)th cause-specific cumulative hazard function with TVC for three TVC models (PE, ED, and CO) can be specified as

\[
H_{f_j}(t|X_{f_i}(t), z_{f_j}) = \int_0^t h_{0j}(u)z_{f_j} \exp \{ m(X_{f_i}(u)) \} \, du
\]

\[
= \begin{cases} 
H_{0j}(t)z_{f_j} & \text{if } t < t_x \text{ (PE,ED,CO)} \\
H_{0j}(t_x)z_{f_j} + \left\{ H_{0j}(t) - H_{0j}(t_x) \right\}z_{f_j} \exp(\beta_j) & \text{if } t \geq t_x \text{ (PE)} \\
H_{0j}(t_x)z_{f_j} + \int_{t_x}^t h_{0j}(u)z_{f_j} \exp \{ \beta_j e^{-\eta_j(u-t_x)} \} \, du & \text{if } t \geq t_x \text{ (ED)} \\
H_{0j}(t_x)z_{f_j} + \int_{t_x}^t h_{0j}(u)z_{f_j} \exp \{ \beta_j e^{-\eta_j(u-t_x)} + \eta_0 \} \, du & \text{if } t \geq t_x \text{ (CO)}
\end{cases}
\]

where \( H_{0j}(t) = \int_0^t h_{0j}(u) \, du \) and numerical integration is required for computing cumulative hazard for ED and CO since no closed form exists.
Web Appendix B:  
Derivation of cause-specific penetrance function with a time-varying covariate

Similarly, with the time varying covariate $x_f(t) = 0$ at $t < t_x$ and 1 at $t \geq t_x$, the marginal cause-specific penetrance function (eq. 6) for the event of interest ($j = 1$) becomes: if $t < t_x$ regardless of TVC models,

$$F_{f_i}(t|X_{f_i}(t)) = \int_0^t h_{01}(u) \left\{ 1 + \frac{H_{01}(u)}{k_1} \right\}^{-k_1-1} \left\{ 1 + \frac{H_{02}(u)}{k_2} \right\}^{-k_2} du,$$

and if $t \geq t_x$, it follows respectively under PE, ED and CO models as:

$$F_{f_i}(t|X_{f_i}(t)) = \int_0^t h_{01}(u) \exp(\beta_1) \times \left[ 1 + \frac{H_{01}(t_x) + \{H_{01}(u) - H_{01}(t_x)\} \exp(\beta_1)}{k_1} \right]^{-k_1-1} \times \left[ 1 + \frac{H_{02}(t_x) + \{H_{02}(u) - H_{02}(t_x)\} \exp(\beta_2)}{k_2} \right]^{-k_2} du, \quad \text{(PE)}$$

$$F_{f_i}(t|X_{f_i}(t)) = \int_0^t h_{01}(u) \exp\left\{ \beta_1 e^{-\eta_1(u-t_x)} \right\} \times \left[ 1 + \frac{H_{01}(t_x) + \int_{t_x}^u h_{01}(s) \exp\left\{ \beta_1 e^{-\eta_1(s-t_x)} \right\} ds}{k_1} \right]^{-k_1-1} \times \left[ 1 + \frac{H_{02}(t_x) + \int_{t_x}^u h_{02}(s) \exp\left\{ \beta_2 e^{-\eta_2(s-t_x)} \right\} ds}{k_2} \right]^{-k_2} du, \quad \text{(ED)}$$

$$F_{f_i}(t|X_{f_i}(t)) = \int_0^t h_{01}(u) \exp\left\{ \beta_1 e^{-\eta_1(u-t_x)} + \eta_0\right\} \times \left[ 1 + \frac{H_{01}(t_x) + \int_{t_x}^u h_{01}(s) \exp\left\{ \beta_1 e^{-\eta_1(s-t_x)} + \eta_0\right\} ds}{k_1} \right]^{-k_1-1} \times \left[ 1 + \frac{H_{02}(t_x) + \int_{t_x}^u h_{02}(s) \exp\left\{ \beta_2 e^{-\eta_2(s-t_x)} + \eta_0\right\} ds}{k_2} \right]^{-k_2} du, \quad \text{(CO)}$$

where $\beta_1$ and $\beta_2$ are the cause-specific TVC effect coefficients, $\eta_1$ and $\eta_2$ are the cause-specific TVC decay rate parameters and $\eta_{01}$ and $\eta_{02}$ are the cause-specific TVC decay convergence parameters for event 1 and 2, respectively.
Web Appendix C: Detailed simulation process

Data were simulated with code modified from the R package ‘FamEvent’ (Choi et al., 2017). Generation of the cause-specific competing risks survival data is based on the algorithm proposed by Beyersmann et al. (2009). Data generation and analyses were performed using R version 3.4.3.

We consider the shared frailty competing risk model with TVC for two competing risks. For the covariates, we include one TIC and one TVC.

1. G: Binary mutation status TIC. If the individual is a mutation carrier, G takes value of 1 otherwise 0. We assume cause specific hazards for both competing events are affected by this variable.

2. X(t): Screening status TVC with the screening time $t_s$. $X(t) = 1$ if $t \geq t_s$ and 0 otherwise. We assume only the cause-specific hazard for event 1 is affected by this variable.

The cause-specific hazards functions for event 1 and event 2 are respectively as follow:

$$h_1(t|X(t), G, z_1) = h_{01}(t) \exp\{\beta_{1\text{gene}}G + m(X(t))\}z_1$$

$$h_2(t|G, z_2) = h_{02}(t) \exp\{\beta_{2\text{gene}}G\}z_2,$$

where $h_{01}(t)$ and $h_{02}(t)$ are the Weibull baseline hazard functions, $z_1$ and $z_2$ are the cause-specific shared frailties, $\beta_{1\text{gene}}$ and $\beta_{2\text{gene}}$ are the mutation status covariate coefficients for event 1 and 2, respectively, and $m(X(t))$ is the effect of the screening TVC, which takes the following form depending on the model:

$$m(X(t)) = \begin{cases} 
0 & \text{if } t < t_s \text{ (PE, ED, CO)} \\
\beta_{\text{screen}} & \text{if } t \geq t_s \text{ (PE)} \\
\beta_{\text{screen}} \exp\{-\eta(t-t_s)\} & \text{if } t \geq t_s \text{ (ED)} \\
\beta_{\text{screen}} \exp\{-\eta(t-t_s)\} + \eta_0 & \text{if } t \geq t_s \text{ (CO)}. 
\end{cases}$$

The algorithm for generating families takes the following three steps based on model (7). Parameters specified in the data generation process, such as the number of siblings for each generation in family pedigree and the current age distribution of
the probands and other family members result in the family structure similar to the real data in the application section.

Step 1: Family structure

1. For each family, we generate a three-generation pedigree. We fix two members in the first generation while we generate 2 to 5 siblings in the second and 0 to 2 siblings in the third generations from a truncated negative binomial distribution.

2. Generate the current age of the proband, \( a_{fp} \) from normal distribution with mean age of 45 and SD of 10. Then we generate the current ages of other family members, \( \{a_{f2}, \ldots, a_{fi}\} \) for individual \( i, i = 2, \ldots, n_f \), from a normal distribution. The current ages of the first generation are generated with the mean age equal to \( a_{fp} + 20 \) with SD of 1.5 years. The current ages of the second generation are generated from mean age equivalent to \( a_{fp} \) with SD of 1.5 years. Finally, for the third generation, their current ages are generated with the mean age subtracted by 20 years from the minimum age of their parents.

3. To generate the screening TVC, we first generate the screening ages \( t_{sf} \) for all members of the family from a normal distribution with mean age of 40 and variance of 2 years. If \( t_{sf,i} > a_{fi} \) we assume this individual does not experience screening.

4. Generate shared frailties \( z_f = \{z_{f1}, z_{f2}\} \) for family \( f \) for two competing events. We assume \( z_{f1} \) and \( z_{f2} \) are independent and marginally follow the gamma distribution with shape parameter \( k_1 \) and the scale parameter \( 1/k_1 \) for event 1 and \( k_2 \) and \( 1/k_2 \) for event 2 respectively.

5. Generate the mutation status variable \( G_{fp} \) for the proband assuming all the probands are the mutation carriers, based on a dominant model with prespecified \( BRCA1 \) mutation allele frequency of 0.0021. Other family members’ mutation statuses are generated conditioning on the proband’s mutation status from a Bernoulli distribution with a probability of success equal to \( P(G_{fi} = 1|G_{fp}) \). This probability depends only on the relationship between the proband and the \( i \)th member of the family by Mendelian inheritance laws.

Step 2: Event times and event types

29
1. Generate $t_{f_i}$ from the overall survival function: Generate $w$ following a uniform on $[0,1]$ and solve for $t_{f_i}$ from $P(T_{f_i} > t_{f_i} | G_{f_i}, t_{s,f_i}, z_f) = w$.

2. Given $t_{f_i}$, we decide the event type $\delta_{f_i}$ among two competing events using the rate of the cause-specific hazards at $t_{f_i}$. Compute $h_1(t_{f_i} | G_{f_i}, t_{s,f_i}, z_f)$, $h_2(t_{f_i} | G_{f_i}, t_{s,f_i}, z_f)$ and $p = \frac{h_1}{h_1 + h_2}$. Run a Bernoulli experiment with the probability of success $p$. If success, then $\delta_{f_i} = 1$ otherwise $\delta_{f_i} = 2$. If $t_{f_i} > a_{f_i}$ we regard this individual as censored and $\delta_{f_i} = 0$. Follow-up duration is defined from age 16 to $a_{f_i}$ if the individual is right censored, otherwise it is from age 16 to $t_{f_i}$.

Step 3: Ascertainment condition for the family

1. After generating the event times and types of the family members, keep the family if it satisfies the condition $t_{f_p} < a_{f_p}$. This condition mimics the population based design of the family studies (Gong and Whittemore, 2003) where probands are affected before their study entry age, $a_{f_p}$.

2. Remove men in the pedigree since the real data only consists of women. Mean pedigree size of 5 leads to the total number of individuals about 2500 when 500 families are generated, which agrees with $BRCA1$ data.

To generate the data, we specify the following parameters:

1. baseline hazard function parameters: $\lambda_1$ and $\rho_1$ for event 1, $\lambda_2$ and $\rho_2$ for event 2

2. parameters involved in TIC: $\beta_{1gene}$ and $\beta_{2gene}$ as genetic effects for each event

3. parameters involved in TVC: $\beta_{screen}$ as a screening effect for event 1, $\eta$ for ED and CO, additional $\eta_0$ for CO

4. familial dependence parameter: $k_1$ and $k_2$ for each event
Web Table 1: Empirical parameter estimates for the competing risks model with a time varying covariate (TVC) under low \( k_1 = 7 \), medium \( k_1 = 3.5 \) and high \( k_1 = 1 \) familial dependence; permanent exposure (PE), exponential decay (ED) or Cox and Oaks (CO) models are considered for TVC. For each scenario, the mean bias, empirical standard error (ESE), average standard error (ASE) and estimated 95\% coverage probability (ECP) are obtained from 500 replicates each with on \( n = 1000 \) families.

| TVC model | True value | \( k_1 = 7, \tau = 0.07 \) | True value | \( k_1 = 3.5, \tau = 0.13 \) | True value | \( k_1 = 1, \tau = 0.33 \) |
|-----------|------------|----------------|------------|----------------|------------|----------------|
| PE log(\( \lambda_1 \)) | -4.83 | -4.83 | -4.83 | -4.83 | -4.83 |
| log(\( \rho_1 \)) | 0.88 | 0.88 | 0.88 | 0.88 | 0.88 |
| log(\( \lambda_2 \)) | -4.96 | -4.96 | -4.96 | -4.96 | -4.96 |
| log(\( \rho_2 \)) | 1.12 | 1.12 | 1.12 | 1.12 | 1.12 |
| \( \beta_{\text{screen}} \) | 0.67 | 0.67 | 0.67 | 0.67 | 0.67 |
| \( \beta_{\text{gene}} \) | 1.95 | 1.95 | 1.95 | 1.95 | 1.95 |
| log(\( k_1 \)) | 1.95 | 1.95 | 1.95 | 1.95 | 1.95 |
| log(\( k_2 \)) | 1.06 | 1.06 | 1.06 | 1.06 | 1.06 |
| ED log(\( \lambda_1 \)) | -4.83 | -4.83 | -4.83 | -4.83 | -4.83 |
| log(\( \rho_1 \)) | 0.83 | 0.83 | 0.83 | 0.83 | 0.83 |
| log(\( \lambda_2 \)) | -4.96 | -4.96 | -4.96 | -4.96 | -4.96 |
| log(\( \rho_2 \)) | 1.08 | 1.08 | 1.08 | 1.08 | 1.08 |
| \( \beta_{\text{screen}} \) | 1.87 | 1.87 | 1.87 | 1.87 | 1.87 |
| \( \beta_{\text{gene}} \) | 1.86 | 1.86 | 1.86 | 1.86 | 1.86 |
| log(\( k_1 \)) | 1.95 | 1.95 | 1.95 | 1.95 | 1.95 |
| log(\( k_2 \)) | 1.06 | 1.06 | 1.06 | 1.06 | 1.06 |
| CO log(\( \lambda_1 \)) | -4.83 | -4.83 | -4.83 | -4.83 | -4.83 |
| log(\( \rho_1 \)) | 0.83 | 0.83 | 0.83 | 0.83 | 0.83 |
| log(\( \lambda_2 \)) | -4.96 | -4.96 | -4.96 | -4.96 | -4.96 |
| log(\( \rho_2 \)) | 1.07 | 1.07 | 1.07 | 1.07 | 1.07 |
| \( \beta_{\text{screen}} \) | 1.52 | 1.52 | 1.52 | 1.52 | 1.52 |
| \( \beta_{\text{gene}} \) | 2.08 | 2.08 | 2.08 | 2.08 | 2.08 |
| log(\( k_1 \)) | 1.95 | 1.95 | 1.95 | 1.95 | 1.95 |
| log(\( k_2 \)) | 1.06 | 1.06 | 1.06 | 1.06 | 1.06 |
| \( \eta_0 \) | 0.21 | 0.21 | 0.21 | 0.21 | 0.21 |

| Bias | ESE | ASE | ECP | Bias | ESE | ASE | ECP | Bias | ESE | ASE | ECP | Bias | ESE | ASE | ECP |
|------|-----|-----|-----|------|-----|-----|-----|------|-----|-----|-----|-----|------|-----|-----|-----|
| PE   | -0.01 | 0.04 | 0.04 | 0.95 | -0.01 | 0.04 | 0.04 | 0.95 | -0.01 | 0.04 | 0.04 | 0.95 | -0.01 | 0.04 | 0.04 | 0.95 |
| ED   | 0.00 | 0.02 | 0.02 | 0.95 | 0.00 | 0.02 | 0.02 | 0.95 | 0.00 | 0.02 | 0.02 | 0.95 | 0.00 | 0.02 | 0.02 | 0.95 |
| CO   | 0.00 | 0.02 | 0.02 | 0.95 | 0.00 | 0.02 | 0.02 | 0.95 | 0.00 | 0.02 | 0.02 | 0.95 | 0.00 | 0.02 | 0.02 | 0.95 |
Web Table 2: Empirical penetrance estimates by age 70 for the competing risks model with a time varying covariate (TVC) under low ($k_1 = 7$), medium ($k_1 = 3.5$) and high ($k_1 = 1$) familial dependence; permanent exposure (PE), exponential decay (ED) or Cox and Oaks (CO) models are considered for TVC; $F_1(70; S, G)$ and $F_2(70; S, G)$ are cause-specific penetrance estimators (%) by age 70 for event 1 and event 2, respectively, given screening (S) and mutation status (G), and screening time at age 35 if $S = 1$. For each scenario, the mean bias, empirical standard error (ESE), average standard error (ASE) and estimated 95% coverage probability (ECP) are obtained from 500 replicates each $n = 1000$ families.

| TVC model | True value | $k_1 = 7$, $\tau = 0.07$ | $k_1 = 3.5$, $\tau = 0.13$ | $k_1 = 1$, $\tau = 0.33$ |
|-----------|------------|--------------------------|--------------------------|--------------------------|
|           | $\lambda$ | Bias | ESE | ASE | ECP | Bias | ESE | ASE | ECP | Bias | ESE | ASE | ECP |
| PE        | $F_1(70; S = 0, G = 0)$ | 12.56 | -0.09 | 0.99 | 0.91 | 0.94 | 12.45 | -0.01 | 1.02 | 0.98 | 0.93 | 11.93 | -0.02 | 0.99 | 1.02 | 0.95 |
|           | $F_1(70; S = 1, G = 0)$ | 21.92 | -0.08 | 1.78 | 1.74 | 0.95 | 21.58 | -0.03 | 1.79 | 1.76 | 0.94 | 20.09 | -0.04 | 1.73 | 1.76 | 0.95 |
|           | $F_1(70; S = 0, G = 1)$ | 56.52 | -0.09 | 2.19 | 2.27 | 0.96 | 54.51 | 0.03 | 2.41 | 2.42 | 0.96 | 46.80 | 0.03 | 2.67 | 2.77 | 0.96 |
|           | $F_1(70; S = 1, G = 1)$ | 75.63 | -0.04 | 2.57 | 2.66 | 0.95 | 72.59 | -0.06 | 2.84 | 2.88 | 0.95 | 61.08 | 0.03 | 3.27 | 3.38 | 0.96 |
|           | $F_2(70; S = 0, G = 0)$ | 4.73 | 0.03 | 0.58 | 0.61 | 0.96 | 4.73 | -0.02 | 0.61 | 0.64 | 0.96 | 4.74 | 0.00 | 0.61 | 0.62 | 0.94 |
|           | $F_2(70; S = 1, G = 0)$ | 4.45 | 0.02 | 0.55 | 0.57 | 0.95 | 4.45 | -0.02 | 0.57 | 0.59 | 0.95 | 4.49 | 0.00 | 0.58 | 0.59 | 0.94 |
|           | $F_2(70; S = 0, G = 1)$ | 9.68 | 0.02 | 0.80 | 0.82 | 0.95 | 9.85 | 0.02 | 0.77 | 0.83 | 0.97 | 10.52 | 0.01 | 0.88 | 0.91 | 0.96 |
|           | $F_2(70; S = 1, G = 1)$ | 7.12 | 0.00 | 0.61 | 0.63 | 0.96 | 7.42 | 0.02 | 0.62 | 0.65 | 0.96 | 8.56 | 0.00 | 0.72 | 0.73 | 0.95 |
| ED        | $F_1(70; S = 0, G = 0)$ | 13.55 | 0.01 | 0.99 | 1.01 | 0.96 | 13.42 | 0.01 | 0.96 | 1.02 | 0.95 | 12.82 | 0.05 | 1.02 | 1.04 | 0.95 |
|           | $F_1(70; S = 1, G = 0)$ | 15.49 | 0.03 | 1.09 | 1.16 | 0.96 | 15.32 | 0.06 | 1.13 | 1.17 | 0.95 | 14.54 | 0.04 | 1.15 | 1.18 | 0.95 |
|           | $F_1(70; S = 0, G = 1)$ | 55.65 | -0.10 | 2.03 | 2.19 | 0.98 | 53.68 | -0.08 | 2.20 | 2.31 | 0.96 | 46.14 | 0.27 | 2.53 | 2.60 | 0.95 |
|           | $F_1(70; S = 1, G = 1)$ | 61.12 | -0.10 | 1.93 | 2.09 | 0.96 | 58.82 | 0.18 | 2.11 | 2.19 | 0.95 | 50.11 | 0.05 | 2.48 | 2.56 | 0.95 |
|           | $F_2(70; S = 0, G = 0)$ | 5.53 | 0.02 | 0.58 | 0.61 | 0.95 | 5.53 | 0.02 | 0.58 | 0.63 | 0.96 | 5.55 | 0.00 | 0.61 | 0.65 | 0.97 |
|           | $F_2(70; S = 1, G = 0)$ | 5.39 | 0.02 | 0.58 | 0.64 | 0.95 | 5.39 | 0.02 | 0.58 | 0.63 | 0.96 | 5.42 | 0.00 | 0.61 | 0.65 | 0.97 |
|           | $F_2(70; S = 0, G = 1)$ | 10.38 | 0.01 | 0.82 | 0.86 | 0.96 | 10.52 | 0.01 | 0.83 | 0.88 | 0.97 | 11.20 | 0.03 | 0.98 | 0.96 | 0.94 |
|           | $F_2(70; S = 1, G = 1)$ | 12.35 | 0.01 | 0.91 | 0.94 | 0.97 | 12.77 | -0.11 | 0.90 | 0.93 | 0.94 | 14.36 | 0.02 | 0.94 | 1.02 | 0.97 |
Web Table 3: Simulation results under a misspecified time varying covariate (TVC): parameter estimates for the competing risks model with a TVC under low ($k_1 = 7$), medium ($k_1 = 3.5$) and high ($k_1 = 1$) familial dependence; permanent exposure (PE), exponential decay (ED) or Cox and Oaks (CO) models are considered as TVC. For each scenario, the three TVC models are fitted and the mean bias, empirical standard error (ESE), average standard error (ASE) and estimated 95% coverage probability (ECP) are obtained from 500 replicates each with $n = 500$ families.

| TVC     | True Model (PE) | Misspecified Model (ED) | Misspecified Model (CO) |
|---------|-----------------|--------------------------|--------------------------|
|         | True $k_1 = 3.5, \tau = 0.13$ | True $k_1 = 3.5, \tau = 0.13$ | True $k_1 = 3.5, \tau = 0.13$ |
| PE      |                 |                          |                          |
| $\log(\lambda_1)$ | -4.83 $0.00 0.06 0.06 0.95$ | -4.83 $0.01 0.06 0.06 0.95$ | -4.83 $0.00 0.06 0.06 0.95$ |
| $\log(\rho_1)$ | 0.88 $0.00 0.03 0.03 0.93$ | 0.88 $0.00 0.03 0.03 0.95$ | 0.88 $0.00 0.03 0.03 0.95$ |
| $\beta_{\text{screen}}$ | 0.67 $0.00 0.10 0.11 0.96$ | 0.67 $0.03 0.13 0.13 0.93$ | 0.67 $-0.30 1.08 0.72 0.54$ |
| $\beta_{\text{gene}}$ | 1.35 $0.01 0.12 0.12 0.96$ | 1.95 $0.00 0.12 0.12 0.95$ | 1.95 $0.00 0.12 0.12 0.94$ |
| $\log(k_1)$ | 1.25 $0.13 0.60 0.48 0.95$ | 1.25 $0.13 0.58 0.51 0.97$ | 1.25 $0.08 0.54 0.50 0.96$ |
| $\log(k_2)$ | 1.06 $0.72 2.20 1.41 0.84$ | 1.06 $0.95 2.49 1.54 0.81$ | 1.06 $0.87 2.27 1.57 0.81$ |
| $\eta$   | -               | -                        | -                        |
| $\eta_0$ | -               | -                        | -                        |
|         |                 |                          |                          |
| ED      |                 |                          |                          |
| $\log(\lambda_1)$ | -4.83 $0.00 0.06 0.06 0.95$ | -4.83 $-0.05 0.07 0.06 0.86$ | -4.83 $0.00 0.06 0.06 0.95$ |
| $\log(\rho_1)$ | 0.83 $0.00 0.03 0.03 0.93$ | 0.83 $-0.02 0.03 0.03 0.87$ | 0.83 $0.00 0.03 0.03 0.95$ |
| $\beta_{\text{screen}}$ | 1.87 $-0.01 0.25 0.25 0.95$ | 1.87 $-1.37 0.13 0.13 0.00$ | 1.87 $0.04 0.26 0.27 0.96$ |
| $\beta_{\text{gene}}$ | 1.86 $0.01 0.11 0.12 0.95$ | 1.86 $0.04 0.13 0.12 0.95$ | 1.86 $0.01 0.11 0.12 0.96$ |
| $\beta_{\text{screen}}$ | 1.22 $0.03 0.22 0.21 0.96$ | 1.22 $0.01 0.23 0.22 0.95$ | 1.22 $0.02 0.20 0.21 0.96$ |
| $\log(k_1)$ | 1.25 $0.08 0.49 0.48 0.97$ | 1.25 $0.21 1.08 0.60 0.86$ | 1.25 $0.11 0.55 0.49 0.96$ |
| $\log(k_2)$ | 1.18 $0.53 1.70 1.26 0.84$ | 1.18 $0.96 1.72 2.12 0.78$ | 1.18 $0.61 1.70 1.46 0.84$ |
| $\eta$   | 0.28 $0.02 0.09 0.09 0.94$ | -               | -                        |
| $\eta_0$ | -               | -                        | -                        |
|         |                 |                          |                          |
| CO      |                 |                          |                          |
| $\log(\lambda_1)$ | -4.83 $0.00 0.06 0.06 0.94$ | -4.83 $-0.02 0.06 0.06 0.94$ | -4.83 $0.03 0.06 0.05 0.96$ |
| $\log(\rho_1)$ | 0.83 $0.00 0.03 0.03 0.93$ | 0.83 $-0.01 0.03 0.03 0.93$ | 0.83 $0.01 0.03 0.03 0.94$ |
| $\log(\lambda_2)$ | -4.96 $0.00 0.07 0.09 0.97$ | -4.96 $0.01 0.09 0.09 0.92$ | -4.96 $-0.01 0.09 0.09 0.95$ |
| $\log(\rho_2)$ | 1.07 $0.00 0.05 0.06 0.97$ | 1.07 $0.01 0.07 0.06 0.92$ | 1.07 $0.00 0.06 0.06 0.94$ |
| $\beta_{\text{screen}}$ | 1.52 $0.04 0.33 0.42 0.94$ | 1.52 $-1.15 0.13 0.12 0.00$ | 1.52 $0.10 0.45 0.42 0.88$ |
| $\beta_{\text{gene}}$ | 2.08 $0.01 0.10 0.12 0.95$ | 2.08 $0.02 0.12 0.12 0.94$ | 2.08 $0.00 0.12 0.12 0.94$ |
| $\beta_{\text{screen}}$ | 1.57 $0.00 0.17 0.21 0.94$ | 1.57 $0.03 0.24 0.21 0.92$ | 1.57 $0.03 0.20 0.21 0.95$ |
| $\log(k_1)$ | 1.25 $0.10 0.39 0.46 0.96$ | 1.25 $0.20 0.82 0.55 0.92$ | 1.25 $0.08 0.48 0.44 0.96$ |
| $\log(k_2)$ | 1.26 $0.35 0.98 1.40 0.90$ | 1.26 $0.60 1.39 1.96 0.80$ | 1.26 $0.52 1.74 1.38 0.86$ |
| $\eta$   | 0.83 $0.01 0.50 0.56 0.91$ | -               | -                        |
| $\eta_0$ | -0.01 $0.12 0.14 0.96$ | -               | -                        |
Web Table 4: Simulation results under a misspecified time varying covariate (TVC): penetrance estimates by age 70 for a competing risks model with a TVC under low \((k_1 = 7)\), medium \((k_1 = 3.5)\) and high \((k_1 = 1)\) familial dependence; permanent exposure (PE), exponential decay (ED) or Cox and Oaks (CO) models are considered for TVC; \(F_1(70; S, G)\) and \(F_2(70; S, G)\) are cause-specific penetrance estimators (%) by age 70 for event 1 and event 2, respectively, given screening (S) and mutation status (G), and screening time at age 35 if \(S = 1\). For each scenario, the three TVC models are fitted and the mean bias, empirical standard error (ESE), average standard error (ASE) and estimated 95% coverage probability (ECP) are obtained from 500 replicates each with \(n = 500\) families.

| True Model (PE) | Misspecified Model (ED) | Misspecified Model (CO) |
|-----------------|-------------------------|-------------------------|
| **TVC**         | \(k_1 = 3.5, \tau = 0.13\) | \(k_1 = 3.5, \tau = 0.13\) | \(k_1 = 3.5, \tau = 0.13\) |
| **PE**          |                         |                         |                         |
| \(F_1(70; S = 0, G = 0)\) | 12.45 0.01 1.33 1.40 0.94 | 12.45 0.21 1.41 1.44 0.95 | 12.45 0.01 1.41 1.42 0.96 |
| \(F_1(70; S = 1, G = 0)\) | 21.58 0.02 2.37 2.48 0.95 | 21.58 -0.35 2.40 2.54 0.95 | 21.58 -0.14 2.56 2.59 0.95 |
| \(F_1(70; S = 0, G = 1)\) | 54.51 0.12 3.39 3.42 0.94 | 54.51 0.39 3.30 3.49 0.96 | 54.51 -0.23 3.30 3.48 0.95 |
| \(F_1(70; S = 1, G = 1)\) | 72.59 0.03 4.08 4.06 0.94 | 72.59 -0.62 3.97 4.26 0.96 | 72.59 -0.57 4.05 4.27 0.96 |
| **ED**          |                         |                         |                         |
| \(F_1(70; S = 0, G = 0)\) | 4.73 -0.08 0.87 0.85 0.93 | 4.73 0.01 0.88 0.87 0.94 | 4.73 -0.05 0.81 0.86 0.95 |
| \(F_1(70; S = 1, G = 0)\) | 4.45 -0.08 0.82 0.80 0.93 | 4.45 0.01 0.83 0.82 0.94 | 4.45 -0.05 0.77 0.81 0.95 |
| \(F_1(70; S = 0, G = 1)\) | 9.85 -0.04 1.16 1.18 0.95 | 9.85 -0.04 1.13 1.18 0.94 | 9.85 0.05 1.05 1.18 0.97 |
| \(F_1(70; S = 1, G = 1)\) | 7.42 -0.04 0.91 0.92 0.95 | 7.42 -0.01 0.87 0.93 0.96 | 7.42 0.04 0.84 0.93 0.96 |
| **CO**          |                         |                         |                         |
| \(F_1(70; S = 0, G = 0)\) | 13.42 -0.02 1.41 1.44 0.94 | 13.42 -0.77 1.63 1.52 0.87 | 13.42 -0.02 1.39 1.49 0.96 |
| \(F_1(70; S = 1, G = 0)\) | 15.32 0.05 1.62 1.66 0.94 | 15.32 3.61 2.38 2.29 0.71 | 15.32 -0.01 1.99 2.01 0.95 |
| \(F_1(70; S = 0, G = 1)\) | 53.68 -0.05 3.03 3.27 0.96 | 53.68 -1.11 3.79 3.54 0.90 | 53.68 0.04 3.29 3.39 0.95 |
| \(F_1(70; S = 1, G = 1)\) | 58.24 0.10 3.26 3.54 0.97 | 58.24 7.37 4.35 4.26 0.61 | 58.24 -0.03 4.20 4.36 0.96 |
| \(F_1(70; S = 0, G = 0)\) | 5.39 -0.07 0.86 0.92 0.95 | 5.39 0.07 0.98 0.97 0.94 | 5.39 -0.05 0.85 0.92 0.95 |
| \(F_1(70; S = 1, G = 0)\) | 5.26 -0.07 0.83 0.89 0.95 | 5.26 -0.03 0.94 0.93 0.94 | 5.26 -0.05 0.83 0.90 0.95 |
| \(F_1(70; S = 0, G = 1)\) | 11.57 0.04 1.29 1.24 0.95 | 11.57 0.20 1.31 1.31 0.95 | 11.57 0.02 1.16 1.25 0.97 |
| \(F_1(70; S = 1, G = 1)\) | 10.22 0.01 1.12 1.12 0.95 | 10.22 -0.53 1.08 1.14 0.90 | 10.22 0.01 1.08 1.16 0.96 |
| **True Model (ED)** | | | |
| **Misspecified Model (PE)** | | | |
| **Misspecified Model (CO)** | | | |
Web Table 5: Comparison of different TVC competing risks models for modelling MS and rrBSO variables in the BRCA1 families from the BCFR.

| TVC for screen | PE for rBSO | ED for rBSO | CO for rBSO |
|----------------|-------------|-------------|-------------|
|                | PE          |            |             |
| βgene          | 2.196       | 2.247       | 2.245       |
| βkronen1       | 0.864       | 0.847       | 0.866       |
| βkronen2       | 0.904       | 0.818       | 0.869       |
| βkronen3       | 0.702       | 0.437       | 0.498       |
| βrrBSO         | -0.754      | -0.947      | -3.961      |
| βgene          | 1.365       | 1.429       | 1.474       |
| βgene          | -0.386      | -0.375      | -0.371      |
| log(ki)        | 1.262       | 1.147       | 1.129       |
| log(kj)        | 0.765       | 0.857       | 0.847       |
| log(rrBSO)     | -4.231      | 2.074       | -0.031      |
| log(rrBSO)     |             |             |             |

| TVC for screen | PE for rBSO | ED for rBSO | CO for rBSO |
|----------------|-------------|-------------|-------------|
|                | AIC          |             |             |
|                 | 19299.28     | 19299.79    | 19297.05    |
|                 | -loglik      |             |             |
|                 | 9634.64      | 9633.90     | 9631.53     |
|                 |               |             |             |
|                 | ED           |             |             |
| βgene          | 2.308        | 2.289       | 2.262       |
| βkronen1       | 3.721        | 3.594       | 3.623       |
| βkronen2       | 3.482        | 3.393       | 3.290       |
| βkronen3       | 3.756        | 3.511       | 3.635       |
| βrrBSO         | -0.566       | -2.775      | -4.223      |
| βrrBSO         | 1.569        | 1.533       | 1.490       |
| βrrBSO         | -0.316       | -0.329      | -0.355      |
| log(rrBSO)     | 1.334        | 1.284       | 1.234       |
| log(rrBSO)     |               |             |             |
|                 | CO           |             |             |
| βgene          | 3.519        | 3.476       | 3.309       |
| βkronen1       | 3.800        | 4.067       | 3.870       |
| βkronen2       | 3.840        | 3.939       | 3.790       |
| βkronen3       | 1.527        | 1.429       | 1.483       |
| βrrBSO         | -0.568       | -4.796      | -4.749      |
| βrrBSO         | -0.347       | -0.376      | -0.351      |
| log(rrBSO)     | 1.252        | 1.203       | 1.149       |
| log(rrBSO)     | 0.876        | 0.927       | 0.832       |
| log(rrBSO)     |               |             |             |

| TVC for screen | PE for rBSO | ED for rBSO | CO for rBSO |
|----------------|-------------|-------------|-------------|
|                | AIC          |             |             |
|                 | 1900.43      | 19086.13    | 19082.56    |
|                 | -loglik      |             |             |
|                 | 9524.22      | 9524.07     | 9521.28     |
|                 |               |             |             |
|                 |               |             |             |
|                 |               |             |             |

* p-value testing for rrBSO effect comparing to PE model based on the likelihood ratio test with df = 2
Web Figure 1: Hazard functions estimated under the different TVC models in the BRCA1 families from the BCFR.
Web Figure 2: Stacked plot of the overall cumulative penetrance estimated from the competing-risks model and showing the contribution of each competing risk event (BC, OC and death) by age in the \textit{BRCA1} families from the BCFR.