Evaluating prognostic accuracy of biomarkers in nested case-control studies

Biostat 892

Xi Meng, Rui Zhang

Overview

- **Nested case–control** (NCC) design is frequently used to conduct biomarker research, but this sampling strategy creates challenges for data analysis.
- To solve this problem, **inverse probability weighted** (IPW) method is proposed for analyzing data from NCC studies.
- Develop estimators for prognostic accuracy measures for a novel biomarker.
- Simulation studies and analysis using Framingham Offspring Study data suggest that the proposed methods perform well in finite samples.

Nested case-control studies

- Nested case-control (NCC): a case-control study nested within a cohort study.
- In a nested-case-control study, all individuals observed to have an event are selected as cases. At each selected case's failure time, a certain size of sample is selected from that case's matched risk set as controls.
- Ideal for predictor variables that are expensive to measure and that can be assessed at the end of the study on subjects who develop the outcome during the study (cases) and on a sample of those who do not (controls).

NCC study in biomarker research

- To enable biomarker research for predicting future events, the biospecimens of the full cohort are often collected at baseline and stored for future studies.
- A standard full cohort design may be infeasible or inefficient for subsequent biomarker studies because assessment of biomarkers can be expensive and labor intensive.
- To overcome such difficulties, the NCC study design is often adopted as a cost-effective cohort sampling strategy.

Approaches to analyze data from NCC

- Conditional logistic regression is typically used to estimate hazard ratio parameters under the proportional hazards (PH) model.
- Stratified sampling of controls in NCC designs has been considered in order to improve efficiency of a simple NCC design.
- Inverse probability weighted (IPW) estimators – contributions of individuals are weighted inversely proportional to their sampling fractions.
  Compared with the traditional partial likelihood–based method, IPW estimators:
  - More efficient as more individuals are included in risk sets.
  - More flexible in estimating functions beyond hazard ratios.

Prognostic biomarker accuracy measures

- Previous developments in NCC study focus on estimating relative risk parameters, which do not fully assess biomarker performance.
- This paper develops estimators for time-dependent prediction accuracy measures based on the idea of IPW.
  - The retrospective accuracy summaries, true-positive fraction (TPF) and false-positive fraction (FPF), evaluated at time t and a cutpoint c:

$$TPF_t(c) = P\{Y > c | T \leq t\}$$

$$FPF_t(c) = P\{Y > c | T \geq t\}$$

  - The prospective accuracy summaries, positive predictive value (PPV) and negative predictive value (NPV)

$$PPV_t(c) = P\{T \leq t | Y > c\}$$

$$NPV_t(c) = P\{T > t | Y \leq c\}$$

General notations

- The observed event time data consist of N iid bivariate vector $\{(X_i, \delta_i), i = 1, ..., N\}$, where $X_i = min(T_i, C_i)$, $\delta_i = I(T_i \leq C_i)$, T_i: event time; C_i: censoring time, N: cohort size.
- Under a NCC design, all cases are selected with their event times denoted as $\{t_1, ..., t_n\}$. At each selected case's failure time t_j, a random sample of size m is selected without replacement from the risk set $R(t_j)$.
  The number of individuals at risk at the selected event time t_j is denoted as $n(t_j) = \sum_i^N I(X_i \geq t_j)$;
- Binary indicator variable V_i, $V_i = 1$ indicates subject i is ever sampled into the NCC subcohort either as a case or as a control.

Estimation with a cohort study

- Using Bayes' theorem, we can rewrite the retrospective prognostic accuracy summaries as:

$$TPF_t(c) = \frac{P(T \leq t, Y > c)}{P(T \leq t)} = \frac{\{1 - F(c)\} - S(t, c)}{1 - S(t)}$$

$$FPF_t(c) = \frac{P(Y > c, T > t)}{P(T > t)} = \frac{S(t, c)}{S(t)}$$

- Similiarly, the prospective accuracy summaries can be written as

$$PPV_t(c) = \frac{\{1 - F(c)\} - S(t, c)}{1 - F(c)}$$

$$NPV_t(c) = \frac{\{S(t) - F(c)\} - S(t, c)}{F(c)}$$

where $S(t, c) = P(T > t, Y > c)$ is the bivariate survival function, $F(c) = P(Y \leq c)$ is the marginal cdf of marker Y.

Estimation with a cohort study (cont.)

- All 4 accuracy summaries involve $F(c)$ and $S(t,c)$.
  One can estimate the marginal distribution of Y and the bivariate survival function as

$$\hat{\mathcal{F}}(y) = \frac{1}{N}\sum_{i=1}^{N} I(Y_i \leq y)$$

$$\hat{\mathcal{S}}(t,c) = \int_c^{\infty} \hat{S}(t|y)d\hat{\mathcal{F}}(y) = \frac{1}{N}\sum_{i=1}^{N} S(t|Y_i)I(Y_i > c)$$

- Plug-in estimators $\widehat{TPR}$, $\widehat{FPR}$, $\widehat{PPV}$, and $\widehat{NPV}$ can be obtained based on the above equations.

Estimation with a cohort study (cont.)

Under a PH model:

$$\lambda(t) = \lambda_0(t)\exp(\beta y)$$

$$S(t|y) = S_0(t)^{\exp(\beta y)} = \exp\{-\Lambda_0(t)\exp(\beta y)\}$$

Thus $S(t|y)$ can be estimated as

$$\hat{S}(t|y) = \exp\{-\hat{\Lambda}_0(t)\exp(\hat{\beta} y)\}$$

- $\hat{\beta}$ is estimated from the partial likelihood
- The Breslow estimator:

$$\hat{\Lambda}_0(t) = \sum_{i=1}^{N} \frac{\delta_i I(X_i \leq t)}{\sum_{j=1}^{N} I(X_j \geq X_i)\exp(\hat{\beta} Y_j)}$$

Sampling probability of an NCC design

- Consider weighing the contributions from the selected observations with IPW weight $\hat{w}_i = V_i/\widehat{p}_i$, where $\widehat{p}_i$ is the probability of the ith subject being selected to the NCC cohort as a case or control based on the sampling scheme.
- The probability subject i is not selected for any of cases

$$\widehat{G}(X_i) = \prod_{j:X_j<X_i} \{1 - \frac{m\delta_j}{n(X_j) - 1}\}$$

- Selection probability

$$\widehat{p}_i = \delta_i + (1 - \delta_i)\{1 - \widehat{G}(X_i)\}$$

Estimating $S(t|y)$

- The log hazard ratio β under the PH model can be obtained by maximizing a weighted partial likelihood with weights accounting for outcome-dependent sampling.

$$\hat{\beta}^w = \text{argmax}_\beta \mathcal{L}(\beta)$$

$$\mathcal{L}(\beta) = \sum_{i=1}^{N} \widehat{w}_i \delta_i \{\beta Y_i - \log \sum_{j=1}^{N} I(X_j \geq X_i) \exp(\beta Y_j)\}$$

- Based on $\hat{\beta}^w$,

$$\widehat{S}^w(t|y) = \exp\{-\widehat{\Lambda}_0^w(t) \exp(\widehat{\beta}^w y)\}$$

where

$$\widehat{\Lambda}_0^w(t) = \sum_{i=1}^{N} \frac{\widehat{w_i} I(X_i \leq t)\delta_i}{\sum_{j=1}^{N} I(X_j \geq X_i)\widehat{w_j} \exp(\widehat{\beta} Y_j)}$$

Estimating $F_Y(y)$ and $S(t,c)$

- Construct IPW estimators for the bivariate survival function of Y and T.

$$\widehat{S}^w(t,c) = \frac{\sum_i^N \widehat{S}^w(t|Y_i)\widehat{w}_i I(Y_i > c)}{\sum_{j=1}^N \widehat{w}_i}$$

- Subsequently, we may estimate $F_Y(c)$ and the marginal survival distribution of T, $S(t)$ as below

$$\widehat{F}_Y^w(c) = 1 - \widehat{S}^w(0,c)$$

$$\widehat{S}^w(t) = \widehat{S}^w(t,-\infty)$$

Estimating accuracy summaries

- Retrospective accuracy summaries can be calculated as

$$\widehat{\text{TPF}}_t^w(c) = \frac{\{1 - \hat{F}^w(c)\} - \hat{S}^w(t,c)}{1 - \hat{S}^w(t)}$$

$$\widehat{\text{FPF}}_t^w(c) = \frac{\hat{S}^w(t,c)}{\hat{S}^w(t)}$$

- The estimators for prospective accuracy summaries are

$$\widehat{\text{PPV}}_t^w(c) = \frac{\{1 - \hat{F}^w(c)\} - \hat{S}^w(t,c)}{1 - \hat{F}^w(c)}$$

$$\widehat{\text{NPV}}_t^w(c) = \frac{\{\hat{S}^w(t) - \hat{F}^w(c)\} - S(t,c)}{\hat{F}^w(c)}$$

Stratified NCC sampling

- A **stratified sampling** of controls based on variables that are correlated with the marker may enhance the power of a simple NCC design.
- In a stratified NCC sampling, at each case's failure time, controls are selected among the risk set and matched to the case based on some covariate Z.
- To incorporate additional matching in the proposed IPW approach, replace $\widehat{w_i}$ in the weighted likelihood with $\widehat{w}_{zi} = V_i/\widehat{p}_{zi}$,

$$\widehat{p}_{zi} = \delta_i + (1-\delta_i)\{1 - \widehat{G}_z(X_i, Z_i)\}$$

$$\widehat{G}_z(X_i, Z_i) = \prod_{j:X_j<X_i, Z_j=Z_i} \{1 - \frac{m\delta_j}{n_z(X_j, Z_j) - 1}\}$$

Simulation Study I: Simple NCC design

- Simulate Y from a standard normal distribution and generate T from a PH model: $\lambda(t) = 0.1\exp(\beta Y)$, where $\beta = \log(3)$, censoring time $C = \min\{2, W\}$, where $W \sim Gamma(2.5, 2)$. Results are based on 2000 simulated data sets.
- Assemble an NCC design with cohort size of $N = 1000$ and $m = 1$ or 3 controls.
- Accuracy summaries of marker Y are evaluated at $t = 1$ and $c_1 = F_Y^{-1}(0.25)$, $c_2 = F_Y^{-1}(0.75)$.

Simulation Study I: measurements

- The accuracy measures are calculated by estimating $S(t|y)$ based on either LB estimators (proposed by *Langholz* and *Borgan*) $S^{LB}(t|y)$ or IPW estimators $S^{w}(t|y)$.
- Standard errors from naive estimators $\widehat{\text{SE}}^{n}$, ignoring correlations among V_i, and standard errors adjusted for correlations $\widehat{\text{SE}}^{a}$, and their corresponding coverage probabilities are calculated for the IPW estimators.
- Relative efficiency is calculated as ratio of empirical variance estimated from full cohort data to the specific estimated variance from NCC samples.

Simulation Study I (cont.)

	True	Ave	SE	$\widehat{\text{SE}}^{n}$	$\widehat{\text{SE}}^{a}$	CP^{n}	CP^{a}	Ave_{LB}	SE_{LB}	RE_{IPW}	RE_{LB}
$m = 1$											
β	1.10	1.12	14.6	13.4	13.3	0.92	0.92	1.13	19.7	0.48	0.22
$\widehat{\text{TPF}}_t(c_1)$	0.95	0.95	1.46	1.35	1.35	0.93	0.93	0.95	1.68	0.42	0.27
$\widehat{\text{FPF}}_t(c_1)$	0.72	0.71	4.49	4.42	4.40	0.95	0.95	0.71	4.58	0.12	0.11
$\widehat{\text{PPV}}_t(c_1)$	0.18	0.18	1.86	2.12	1.83	0.97	0.94	0.19	2.73	0.71	0.32
$\widehat{\text{NPV}}_t(c_1)$	0.97	0.97	0.73	0.71	0.68	0.95	0.93	0.97	0.79	0.68	0.50
$\widehat{\text{TPF}}_t(c_2)$	0.60	0.60	4.81	4.76	4.75	0.94	0.94	0.60	5.80	0.47	0.31
$\widehat{\text{FPF}}_t(c_2)$	0.19	0.19	3.63	3.51	3.50	0.94	0.93	0.19	3.87	0.14	0.12
$\widehat{\text{PPV}}_t(c_2)$	0.35	0.35	4.40	4.43	4.15	0.96	0.94	0.36	6.99	0.48	0.17
$\widehat{\text{NPV}}_t(c_2)$	0.92	0.92	1.06	1.17	1.04	0.98	0.94	0.92	1.10	0.80	0.71
$m = 3$											
β	1.10	1.11	10.9	10.8	10.8	0.94	0.94	1.11	13.4	0.70	0.45
$\widehat{\text{TPF}}_t(c_1)$	0.95	0.95	1.04	1.04	1.04	0.96	0.96	0.95	1.18	0.67	0.52
$\widehat{\text{FPF}}_t(c_1)$	0.72	0.71	2.74	2.74	2.74	0.95	0.95	0.71	2.76	0.31	0.30
$\widehat{\text{PPV}}_t(c_1)$	0.18	0.18	1.63	1.72	1.64	0.96	0.95	0.19	1.83	0.85	0.68
$\widehat{\text{NPV}}_t(c_1)$	0.97	0.97	0.59	0.61	0.6	0.96	0.95	0.97	0.65	0.78	0.66
$\widehat{\text{TPF}}_t(c_2)$	0.60	0.60	3.94	3.86	3.85	0.95	0.95	0.60	4.30	0.73	0.59
$\widehat{\text{FPF}}_t(c_2)$	0.19	0.19	2.24	2.22	2.22	0.94	0.94	0.19	2.36	0.35	0.31
$\widehat{\text{PPV}}_t(c_2)$	0.35	0.35	3.37	3.39	3.31	0.95	0.95	0.35	4.25	0.76	0.46
$\widehat{\text{NPV}}_t(c_2)$	0.92	0.92	0.97	1.01	0.97	0.96	0.95	0.92	0.98	0.86	0.84

Simulation Study I (cont.)

- Both IPW estimator and LB estimator are unbiased. However, the IPW estimators in general are more efficient than the LB estimators.
- Adjusted standard errors performed well, with coverage percentage close to 95%.
- The naive standard errors in most of the cases are quite close to their adjusted counterparts, indicating that correlations are weak among observations. However, they do appear to be more conservative when $m = 1$.

Simulation Study II: Complex NCC design

- Generated from the same model except for Y and C are more complicated.
  - Y is generated from a mixture of normal distributions: $Y = \mathcal{B}W_1 + (1-\mathcal{B})W_2$, where $\mathcal{B} \sim$ Bernoulli(0.9), $W_1 \sim N(0, 0.5)$, $W_2 \sim N(0.3, 0.1)$.
  - To introduce biomarker-dependent censoring, $C \sim$ Uniform$(0.5, 1.5)$ if $\mathcal{B} = 1$, $C = \exp(Z/-3Y)$ with $Z \sim N(0,1)$ if $\mathcal{B} = 0$
- Both set of estimators are unbiased. The adjusted variance estimators also work well. However, the naive variance estimators for PPVt(c) and NPVt(c) again could have inflated values when $m = 1$.

Simulation Study II (cont.)

	True	Ave	SE	$\widehat{\text{SE}}^{n}$	$\widehat{\text{SE}}^{a}$	CP^{n}	CP^{a}	Ave_{LB}	SE_{LB}	RE_{IPW}	RE_{LB}
$m = 1$											
β	1.10	1.11	20.0	19.3	19.3	0.94	0.94	1.11	22.7	0.50	0.39
$\widehat{\text{TPF}}_t(c_1)$	0.97	0.97	0.93	0.88	0.88	0.93	0.93	0.97	0.94	0.44	0.44
$\widehat{\text{FPF}}_t(c_1)$	0.89	0.89	2.27	2.21	2.21	0.93	0.93	0.89	2.30	0.23	0.22
$\widehat{\text{PPV}}_t(c_1)$	0.26	0.26	1.78	2.10	1.81	0.98	0.95	0.26	1.81	0.82	0.79
$\widehat{\text{NPV}}_t(c_1)$	0.91	0.91	1.85	1.91	1.84	0.95	0.94	0.91	1.97	0.64	0.56
$\widehat{\text{TPF}}_t(c_2)$	0.18	0.18	2.61	2.64	2.64	0.95	0.95	0.18	2.91	0.62	0.50
$\widehat{\text{FPF}}_t(c_2)$	0.06	0.06	1.49	1.46	1.46	0.93	0.93	0.06	1.50	0.27	0.26
$\widehat{\text{PPV}}_t(c_2)$	0.48	0.49	5.75	5.75	5.58	0.94	0.93	0.49	6.68	0.50	0.37
$\widehat{\text{NPV}}_t(c_2)$	0.78	0.78	1.63	1.91	1.65	0.98	0.95	0.78	1.58	0.82	0.88
$m = 3$											
β	1.10	1.10	15.9	15.5	15.5	0.94	0.94	1.11	17.9	0.81	0.64
$\widehat{\text{TPF}}_t(c_1)$	0.97	0.97	0.70	0.68	0.68	0.93	0.93	0.97	0.74	0.76	0.68
$\widehat{\text{FPF}}_t(c_1)$	0.89	0.89	1.48	1.46	1.46	0.94	0.94	0.89	1.49	0.54	0.54
$\widehat{\text{PPV}}_t(c_1)$	0.26	0.26	1.67	1.76	1.70	0.95	0.95	0.26	1.67	0.91	0.91
$\widehat{\text{NPV}}_t(c_1)$	0.91	0.91	1.62	1.60	1.58	0.94	0.94	0.91	1.71	0.86	0.77
$\widehat{\text{TPF}}_t(c_2)$	0.18	0.18	2.22	2.20	2.20	0.94	0.94	0.18	2.34	0.85	0.77
$\widehat{\text{FPF}}_t(c_2)$	0.06	0.06	0.97	0.99	0.99	0.95	0.95	0.06	1.00	0.61	0.58
$\widehat{\text{PPV}}_t(c_2)$	0.48	0.48	4.51	4.50	4.46	0.94	0.94	0.48	5.12	0.80	0.62
$\widehat{\text{NPV}}_t(c_2)$	0.78	0.78	1.53	1.61	1.56	0.96	0.95	0.78	1.49	0.91	0.96

Simulation Study III: Stratified NCC design

- Controls are also matched with cases on a binary covariate Z.
- With a cohort of size 2000, first generate $Y \sim N(0,1)$, then generate a binary Z s.t. $Z = 1$ if $Y > 0$. T and C are using the same model as the first simulation.
- IPW estimators under stratified NCC design perform well. The naive variance, without considering the correlation due to finite sampling, often lead to overestimated variances.

Simulation Study III (cont.)

	True	Ave	SE	$\widehat{\text{SE}}^{n}$	$\widehat{\text{SE}}^{a}$	CP^{n}	CP^{a}	Ave_{bl}	SE_{bl}	RE_{ipw}	RE_{bl}
$m = 1$											
β	1.10	1.11	8.60	9.30	8.49	0.96	0.94	1.11	17.07	0.60	0.15
$\widehat{\text{TPF}}_t(c_1)$	0.95	0.95	1.23	1.24	1.22	0.93	0.92	0.95	1.56	0.26	0.16
$\widehat{\text{FPF}}_t(c_1)$	0.72	0.72	4.79	5.07	4.76	0.95	0.94	0.72	4.93	0.05	0.05
$\widehat{\text{PPV}}_t(c_1)$	0.18	0.19	1.47	1.74	1.45	0.98	0.95	0.19	2.75	0.51	0.14
$\widehat{\text{NPV}}_t(c_1)$	0.97	0.97	0.50	0.56	0.50	0.96	0.94	0.97	0.64	0.64	0.39
$\widehat{\text{TPF}}_t(c_2)$	0.60	0.60	3.23	3.16	3.15	0.94	0.94	0.60	4.51	0.56	0.27
$\widehat{\text{FPF}}_t(c_2)$	0.19	0.19	2.00	2.55	2.01	0.99	0.95	0.19	2.48	0.20	0.13
$\widehat{\text{PPV}}_t(c_2)$	0.35	0.35	2.56	2.71	2.48	0.97	0.95	0.36	6.29	0.64	0.10
$\widehat{\text{NPV}}_t(c_2)$	0.92	0.92	0.73	1.01	0.75	0.99	0.96	0.92	0.82	0.75	0.63
$m = 3$											
β	1.10	1.10	7.10	7.36	7.07	0.95	0.94	1.11	12.15	0.81	0.27
$\widehat{\text{TPF}}_t(c_1)$	0.95	0.95	0.90	0.88	0.87	0.93	0.93	0.95	1.17	0.49	0.27
$\widehat{\text{FPF}}_t(c_1)$	0.72	0.72	3.06	3.10	2.94	0.95	0.94	0.72	3.14	0.13	0.12
$\widehat{\text{PPV}}_t(c_1)$	0.18	0.18	1.23	1.32	1.22	0.97	0.95	0.19	2.04	0.75	0.27
$\widehat{\text{NPV}}_t(c_1)$	0.97	0.97	0.43	0.45	0.43	0.96	0.94	0.97	0.51	0.78	0.56
$\widehat{\text{TPF}}_t(c_2)$	0.60	0.60	2.60	2.62	2.61	0.95	0.95	0.60	3.45	0.79	0.45
$\widehat{\text{FPF}}_t(c_2)$	0.19	0.19	1.33	1.60	1.36	0.98	0.95	0.19	1.63	0.44	0.30
$\widehat{\text{PPV}}_t(c_2)$	0.35	0.35	2.15	2.21	2.17	0.95	0.95	0.35	4.52	0.84	0.19
$\widehat{\text{NPV}}_t(c_2)$	0.92	0.92	0.69	0.79	0.69	0.97	0.95	0.92	0.75	0.89	0.75

Example of Framingham Offspring Study

- Evaluating the accuracy of an inflammation marker, C-reactive protein (CRP), for predicting the risk of cardiovascular disease (CVD) using the Framingham Offspring data.
- The Framingham Offspring Study was initiated in 1971 with a cohort of 5124 participants. Here consider 3289 participants with CRP measurements at the second examination, consider outcome as time from examination date to first major CVD event.
- From the full cohort data, further assemble NCC subcohorts with 1 or 3 controls who were selected either (i) without additional matching or (ii) matched to their corresponding cases based on gender and age groups.
- Consider low and high thresholds, set as the 25th or 75th percentile of the CRP levels in the full cohort.

Example of Framingham Offspring Study (cont.)

	Full cohort		Not stratified				Stratified			
			$m=1$		$m=3$		$m=1$		$m=3$	
	Ave	SE	Ave	SE	Ave	SE	Ave	SE	Ave	SE
$\widehat{\beta}$	0.483	0.049	0.510	0.133	0.455	0.089	0.508	0.110	0.492	0.080
$c=\widehat{F}_Y^{-1}(0.25)$										
$\widehat{\mathrm{TPF}}_t(c)$	0.900	0.011	0.913	0.027	0.891	0.022	0.906	0.025	0.904	0.017
$\widehat{\mathrm{FPF}}_t(c)$	0.749	0.008	0.748	0.049	0.739	0.030	0.760	0.049	0.787	0.027
$\widehat{\mathrm{PPV}}_t(c)$	0.044	0.004	0.053	0.008	0.053	0.007	0.053	0.006	0.051	0.005
$\widehat{\mathrm{NPV}}_t(c)$	0.985	0.002	0.984	0.005	0.981	0.004	0.982	0.004	0.979	0.004
$c=\widehat{F}_Y^{-1}(0.75)$										
$\widehat{\mathrm{TPF}}_t(c)$	0.471	0.025	0.473	0.057	0.448	0.044	0.482	0.043	0.462	0.033
$\widehat{\mathrm{FPF}}_t(c)$	0.242	0.008	0.225	0.045	0.228	0.028	0.243	0.041	0.272	0.029
$\widehat{\mathrm{PPV}}_t(c)$	0.070	0.007	0.088	0.018	0.083	0.014	0.085	0.013	0.074	0.009
$\widehat{\mathrm{NPV}}_t(c)$	0.974	0.003	0.970	0.005	0.968	0.005	0.969	0.004	0.966	0.004

Example of Framingham Offspring Study (cont.)

- The estimates from NCC studies are quite close to the results using the full cohort data, suggesting that time-dependent accuracy summaries can be reliably estimated from the NCC data with the proposed methods.
- Matching on age and gender improves the efficiency of most estimates slightly, and including more controls results in more precise inference.
- Compared to a full cohort analysis, the standard errors for TPF and FPF, although relatively larger using NCC samples, are still sufficiently precise for making decisions.

Remarks

- The NCC study is a useful design option in the field of biomarker research. By matching controls to cases' failure times and other potential confounding factors, the accuracy of biomarkers can be evaluated more efficiently.
- However, matching generates complex data that can be difficult to analyze. This paper proposed estimators for accuracy measures under an NCC design based on the IPW approach.
- This approach yields more efficient estimators for those accuracy summaries compared to estimators derived from a partial likelihood.
- Compared with a nonparametric MLE-based approach, this approach is very simple to implement and is robust to marker-dependent censoring.
- Extension of this approach to models with multiple covariates and an estimation of covariate-specific accuracy is straightforward.