BIOMETRICS 64, 1–7 DOI: 10.1111/j.1541-0420.2005.00454.x
December 2008

Real-time individual predictions of prostate cancer recurrence using joint models: Online supplementary material

Jeremy M. G. Taylor[1,*], Yongseok Park[1], Donna P. Ankerst[2,3], Cecile Proust-Lima[4], Scott Williams[5], Larry Kestin[6], Kyoungwha Bae[7], Tom Pickles[8], Howard Sandler[9]

[1] Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109

[2] Zentrum Mathematik, Technische Universitaet Muenchen, Munich, Germany

[3] Dept. of Epidemiology/Biostatistics and Urology, University of Texas Health Science Center at San Antonio, San Antonio, Texas

[4] INSERM, U897, Epidemiology and Biostatistics Research Center, 33076 Bordeaux, France

[5] Division of Radiation Oncology, Peter MacCallum Cancer Centre, Melbourne, Victoria, Australia

[6] Michigan Healthcare Professionals/21st Century Oncology, Farmington Hills, MI 48334

[7] Alcon Laboratories Novartis Company, Fort Worth, TX

[8] British Columbia Cancer Agency, Vancouver, BC, Canada

[9] Department of Radiation Oncology, Cedars-Sinai Medical Center, Los Angeles, CA

**email:* jmgt@umich.edu

1. Online Supplementary Material

1.1 *MCMC algorithm for the training data*

- The form of the longitudinal model is:

$$\mathrm{Y}_i(t) = \mathrm{EY}_i(t) + \epsilon_{it} = \boldsymbol{U}(t)^T \boldsymbol{\beta}_i + \epsilon_{it} = \beta_{0i} + \beta_{1i}\{(1+t)^{-1.5} - 1\} + \beta_{2i} t + \epsilon_{it}, \qquad (1)$$

  where $\epsilon_{it} \sim T(0, \sigma^2, 5)$, $\boldsymbol{U}(t)^T = (1, (1+t)^{-1.5} - 1, t)$ and $\boldsymbol{\beta}_i^T \sim N_3(\boldsymbol{x}_i^T \boldsymbol{\alpha}, \boldsymbol{\Sigma}_{3\times3})$.

- The longitudinal model covariates matrix $\boldsymbol{x}_i = (\boldsymbol{x}_{0i}, \boldsymbol{x}_{1i}, \boldsymbol{x}_{2i})^T$ is composed of:

(1) $\boldsymbol{x}_{0i} = (1, \mathrm{PSA}_i, 0, \ldots, 0)^T$, corresponding to parameters (α_1, α_2);

(2) $\boldsymbol{x}_{1i} = (0, 0, 1, \mathrm{T2}_i, \mathrm{T34}_i, \mathrm{PSA}_i, 0, \ldots,)^T$, corresponding to parameters $(\alpha_3, \alpha_4, \alpha_5, \alpha_6)$

(3) $\boldsymbol{x}_{2i} = (0, \ldots, 0, 1, \mathrm{T2}_i, \mathrm{T34}_i, \mathrm{G7}_i, \mathrm{G8U}_i, \mathrm{PSA}_i)^T$, corresponding to parameters $(\alpha_7, \ldots, \alpha_{12})$;

(4) Where the notation for the baseline variables is:

  - $\mathrm{T2}_i = I(\text{T-stage}_i = 2)$,
  - $\mathrm{T34}_i = I(\text{T-stage}_i \geqslant 3)$,
  - $\mathrm{G7}_i = I(\mathrm{Gleason}_i = 7)$,
  - $\mathrm{G8U}_i = I(\mathrm{Gleason}_i \geqslant 8)$.
  - $\mathrm{PSA}_i = \log(\mathrm{basePSA}_i + 0.1)$,

- The fixed effect parameters in the longitudinal model are:

(1) $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_{12})$: 12 parameters;

(2) σ^2: 1 parameter;

(3) $\boldsymbol{\Sigma}$: 3×3 covariance matrix with 6 parameters.

- The form of the survival model is:

$$\lambda_i(t) = \lambda_0(t) \exp\{\boldsymbol{Z}_i(t)^T \boldsymbol{\theta}\}, \qquad (2)$$

  where

$$\boldsymbol{Z}_i(t)^T \boldsymbol{\theta} = [\boldsymbol{W}_i^T \boldsymbol{\theta}_0 + \theta_6 \mathrm{logit}^{-1}(\mathrm{EY}(t)_i - 0.7)/0.45 + \theta_7 \mathrm{EY}_i'(t) + \theta_8 I(t \geqslant H_i)],$$

$\boldsymbol{\theta}_0 = \{\theta_1, \ldots, \theta_5\}^T$, $W_i = \{\mathrm{T2}_i, \mathrm{T34}_i, \mathrm{PSA}_i, \mathrm{G7}_i, \mathrm{G8U}_i\}^T$, H_i is salvage hormone therapy time, and the baseline hazard is given by

- $\lambda_0(t) = \sum_{j=1}^{5} \lambda_{0j} I(B_j < t \leqslant B_{j+1})$, where $B = (0.95, 2, 3, 5, 7, 30)$,
- $\lambda_0(t) = 0$ for $t \leqslant 0.95$.

- The 13 parameters in the survival model are:
  - $\boldsymbol{\theta} = (\theta_1, ..., \theta_8)$, 8 parameters
  - $\boldsymbol{\lambda}_0 = (\lambda_{01}, ..., \lambda_{05})$, 5 parameters
  - The joint likelihood of the observations conditional on the fixed and random effects is given by:

$$\begin{aligned} L \propto & \prod_i^n \lambda_i(R_i)^{\delta_i} \exp\left\{-\sum_i \int_0^{R_i} \lambda_i(t)dt\right\} \\ & \times \sigma^{-m} \exp\left[-\sum_i \sum_j \frac{\{Y_{ij} - \boldsymbol{U}_{ij}^T \boldsymbol{\beta}_i\}^2 s_{ij}}{2\sigma^2}\right] \\ & \times |\Sigma|^{-n/2} \exp\left\{-\sum_i (\boldsymbol{\beta}_i - \boldsymbol{x}_i^T \alpha)^T \Sigma^{-1} (\boldsymbol{\beta}_i - \boldsymbol{x}_i^T \alpha)\right\}, \end{aligned}$$

    where $Y_{ij} = Y_i(t_{ij}), U_{ij} = U(t_{ij})$; R_i is the observed time of clinical recurrence or censoring and δ_i is event indicator for subject i; t_{ij} is PSA longitudinal data measurement time for subject i; s_{ij} is the scaling factor for the T distribution; m is the total observed number of longitudinal PSA measurements; and n is the total number of subjects.
  - The prior distributions are:

$$\begin{aligned} \lambda_{0j} & \sim \mathrm{gamma}(0.01, 0.01), j = 1, \cdots, 5; \\ \sigma^{-2} & \propto 1; \\ \boldsymbol{\Sigma}^{-1} & \sim \mathrm{Wishart}\left(\boldsymbol{\Sigma}_p, 13\right), \text{where } \boldsymbol{\Sigma}_p^{-1} = 0.5\boldsymbol{I}_{3\times 3}; \\ \boldsymbol{\alpha} & \sim N_{12}\left(\boldsymbol{\alpha}_p, \boldsymbol{\Sigma}_\alpha\right), \text{where } \boldsymbol{\alpha}_p = (0.01, \cdots, 0.01)^T, \boldsymbol{\Sigma}_\alpha^{-1} = 0.2 I_{12\times 12}; \\ s_{ij} & \sim \mathrm{gamma}\left(\nu/2, \nu/2\right), \text{ where } \nu = 5; \end{aligned}$$

– Conditional distributions used in the MCMC algorithm :

$$\lambda_{0j} \sim \text{gamma}\Big(0.01 + \sum_{i=1}^{n} I(B_j < R_i \leqslant B_{j+1})\delta_i,$$

$$0.01 + \sum_i \int_{B_j}^{\min(B_{j+1},R_i)} I(R_i > B_j)\exp[\boldsymbol{Z}_i(t)^T\boldsymbol{\theta}]dt\Big), j = 1, \cdots, 5;$$

$$\sigma^{-2} \sim \text{gamma}\left\{\frac{m}{2}, \frac{1}{2}\sum_i\sum_j (Y_{ij} - \boldsymbol{W}_{ij}^T\boldsymbol{\beta}_i)^2 s_{ij}\right\};$$

$$s_{ij} \sim \text{gamma}\left\{\frac{\nu+1}{2}, \frac{\nu}{2} + \frac{(Y_{ij} - \boldsymbol{W}_{ij}^T\boldsymbol{\beta}_i)^2}{2\sigma^2}\right\};$$

$$\boldsymbol{\Sigma}^{-1} \sim \text{Wishart}\left(\left[\sum_i\{(\boldsymbol{\beta}_i - \boldsymbol{x}_i^T\alpha)(\boldsymbol{\beta}_i - \boldsymbol{x}_i^T\boldsymbol{\alpha})^T + \boldsymbol{\Sigma}_p^{-1}\}\right]^{-1}, n+13\right);$$

$$\boldsymbol{\alpha} \sim N_{12}\left[\boldsymbol{\alpha}_p + \sum_i(\boldsymbol{x}_i^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\beta}_i), \{\boldsymbol{\Sigma}_\alpha^{-1} + \sum_i(\boldsymbol{x}_i^T\boldsymbol{\Sigma}^{-1}\boldsymbol{x}_i)\}^{-1}\right],$$

$$\boldsymbol{\alpha}_p = (0.01, \cdots, 0.01)^T, \boldsymbol{\Sigma}_\alpha^{-1} = 0.2I_{12\times12};$$

– MCMC algorithm to draw $\boldsymbol{\beta}_i$ using the Metropolis-Hastings method: draw $\boldsymbol{\beta}_i^{\text{new}}$ from the conditional distribution:

$$\boldsymbol{\beta}_i \sim N_3(\boldsymbol{\mu}, \boldsymbol{V});$$

$$\boldsymbol{\mu} = \boldsymbol{V}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{x}_i^T\boldsymbol{\alpha} + \sum_{j=1} Y_{ij}\boldsymbol{W}_{ij}s_{ij}\sigma^{-2}\right)$$

$$\boldsymbol{V} = \left(\boldsymbol{\Sigma}^{-1} + \sum_{j=1}\boldsymbol{W}_{ij}^T\boldsymbol{W}_{ij}s_{ij}\sigma^{-2}\right)^{-1}.$$

Let

$$r = \exp\left\{\sum_i \int_0^{R_i} \lambda_i(t|\boldsymbol{\theta}, \boldsymbol{\beta}_i)dt - \int_0^{R_i} \lambda_i(t|\boldsymbol{\theta}, \boldsymbol{\beta}_i^{(new)})dt\right\}.$$

If $r \geqslant 1$, set $\boldsymbol{\beta}_i = \boldsymbol{\beta}_i^{(new)}$. Otherwise set $\boldsymbol{\beta}_i = \boldsymbol{\beta}_i^{(new)}$ with probability r.

– MCMC algorithm to draw $\boldsymbol{\theta}$ using Metropolis-Hastings method: draw $\boldsymbol{\theta}^{(\text{new})}$ from predefined $N_8(\boldsymbol{\mu}_8, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})$. The specific values for $\boldsymbol{\mu}_8$ and $\boldsymbol{\Sigma}_\theta$ were based on a preliminary Cox model analyses of the University of Michigan data with imputed values of PSA and slope of PSA.

$\boldsymbol{\mu}_8 = (1.08, 1.33, -0.41, -1.35, 1.03, 1.49, 4.65, 0.42)^T$

$$\boldsymbol{\Sigma}_\theta = \begin{pmatrix} 0.0981 & 0.0887 & -0.0017 & 0.0027 & -0.0073 & 0.0006 & -0.0022 & -0.0041 \\ & 0.1175 & -0.0056 & 0.0049 & -0.0092 & -0.0086 & -0.0099 & -0.0032 \\ & & 0.0113 & 0.0009 & -0.0020 & -0.0042 & -0.0116 & 0.0016 \\ & & & 0.0631 & -0.0042 & -0.0141 & -0.0109 & -0.0092 \\ & & & & 0.0362 & 0.0246 & 0.0083 & -0.0047 \\ & & & & & 0.0793 & 0.0177 & -0.0079 \\ & & & & & & 0.1597 & -0.0374 \\ & & & & & & & 0.0283 \end{pmatrix}$$

Let

$$r = \exp\Bigg[\left\{\sum_i \int_0^{R_i} \lambda_i(t|\boldsymbol{\theta}, \boldsymbol{\beta}_i)dt - \int_0^{R_i} \lambda_i(t|\boldsymbol{\theta}^{(new)}, \boldsymbol{\beta}_i)dt\right\} \\ + (\boldsymbol{\theta} - \boldsymbol{\mu}_8)^T \boldsymbol{\Sigma}_\theta^{-1}(\boldsymbol{\theta} - \boldsymbol{\mu}_8) - (\boldsymbol{\theta}^{(\text{new})} - \boldsymbol{\mu}_8)^T \boldsymbol{\Sigma}_\theta^{-1}(\boldsymbol{\theta}^{(\text{new})} - \boldsymbol{\mu}_8)\Bigg].$$

If $r \geqslant 1$, set $\boldsymbol{\theta} = \boldsymbol{\theta}^{(new)}$. Otherwise set $\boldsymbol{\theta} = \boldsymbol{\theta}^{(new)}$ with probability r.

- The MCMC algorithm was run for 500,000 iterations after a 100,000 iteration burn-in period. The proposal densities for the Metropolis-Hastings draws gave approximately a 2.33% acceptance rate for survival model parameters and 97.5% for random effects ($\boldsymbol{\beta}_i$). The convergence was assessed graphically. 1,000 draws from the converged algorithm were saved, consisting of every five hundredth draw from the final 500,000 draws.

1.2 *Estimates of fixed effects from training data*

[Table 1 about here.]

1.3 *Quick MCMC algorithm for the testing data*

1.3.1 *Notation and description of the necessary elements.*

- $\boldsymbol{Y} = (Y_1, \cdots, Y_m)^T$: longitudinal PSA data of new patient at times $t_1, \ldots, t_m$;

- $\boldsymbol{U} = (U_1, \ldots, U_m)$, where $U_i = (1, (1+t_i)^{-1.5} - 1, t_i)$;
- $\boldsymbol{W}$: baseline covariates of new patient;
- $\boldsymbol{x}$: 12 by 3 matrix of intercepts and baseline covariates, derived from $\boldsymbol{W}$;
- c: truncation time from which the prediction will be made (usually today);
- R: random variable of time to recurrence for the new patient;
- $\boldsymbol{\alpha}, \boldsymbol{\theta}$: fixed effect estimates obtained from the training data;
- $\boldsymbol{\beta}_N$: random effect of new patient to be estimated;
- $(s_1, \cdots, s_m)$: scaling factors to give T distribution for measurement error;
- $Z(t)$: time-dependent covariate vector in hazard model, derived from $\boldsymbol{W}$ and $\boldsymbol{\beta}_N$.

Assumed distributions:

(1) $(\boldsymbol{\alpha}, \boldsymbol{\theta}, \boldsymbol{\Sigma}, \sigma^2, \boldsymbol{\lambda}_0 | \text{Training Data})$: posterior draws of parameters based on analysis of the training data, which is independent of the new patient;

(2) $\boldsymbol{\beta}_N \sim N(\boldsymbol{x}^T \boldsymbol{\alpha}, \boldsymbol{\Sigma})$: distribution of random effects (multivariate normal).

(3) $Y_j \sim \mathrm{T}(\boldsymbol{U}_j \boldsymbol{\beta}_N, \sigma^2, 5)$: distribution of longitudinal data (T_5 distribution);

(4) $R | \boldsymbol{W}, \boldsymbol{\beta}_N, \boldsymbol{\theta}, \boldsymbol{\lambda}_0, R > c$: residual time to recurrence distribution;

The quantities to be estimated

- $P(R > c + t | \boldsymbol{W}, \boldsymbol{Y}, R > c, \text{Training data})$;
- $\mathrm{EY}(t | \boldsymbol{x}, \boldsymbol{Y}, R > c, \text{Training data})$.

The likelihood for the new patient

$$f(\boldsymbol{\beta}_N, s_1, \ldots, s_m) \propto \prod_{j=1}^{m} \left[s_j^{1/2} \exp\left\{ -\frac{(Y_j - \boldsymbol{U}_j \boldsymbol{\beta}_N)^2 s_j}{2\sigma^2} \right\} \right] s_j^{\nu/2-1} \exp\Big\{ -\frac{\nu s_j}{2} \Big\}$$
$$\exp\left\{ -\frac{1}{2} (\boldsymbol{\beta}_N - \boldsymbol{x}^T \boldsymbol{\alpha})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\beta}_N - \boldsymbol{x}^T \boldsymbol{\alpha}) \right\} \exp\left[-\int_0^c \lambda_0(t) \exp\{ \boldsymbol{Z}(t)^T \boldsymbol{\theta} \} dt \right],$$

where $\nu = 5$.

Conditional distributions in the MCMC algorithm:

$$s_j \sim \text{gamma}\Big(\frac{\nu+1}{2}, \frac{(Y_j - \boldsymbol{U}_j \boldsymbol{\beta}_N)^2}{2\sigma^2} + \frac{\nu}{2}\Big)$$

Metropolis-Hastings draw for $\boldsymbol{\beta}_N$: The proposal distribution for $\boldsymbol{\beta}_N$ is trivariate normal as described below.

1.3.2 *Algorithm.*

(1) Draw $(\boldsymbol{\alpha}, \boldsymbol{\theta}, \sigma^2, \boldsymbol{\Sigma}, \boldsymbol{\lambda}_0)$ from posterior distribution derived from training data set;

(2) Set initial values. Let $s_j = 1, j = 1, \cdots, m$ and $\boldsymbol{\beta}_N = \boldsymbol{x}^T\boldsymbol{\alpha}$;

(3) Draw $\boldsymbol{\beta}_N^{(new)}$ from multivariate normal $\mathcal{N}_3(\boldsymbol{\mu}, \boldsymbol{V})$ proposal distribution, where

$$\boldsymbol{\mu} = \boldsymbol{V}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{x}^T\boldsymbol{\alpha} + \sum_{j=1}^{m} Y_j \boldsymbol{U}_j s_j \sigma^{-2}\right)$$

$$\boldsymbol{V} = \left(\boldsymbol{\Sigma}^{-1} + \sum_{j=1}^{m} \boldsymbol{U}_j^T \boldsymbol{U}_j s_j \sigma^{-2}\right)^{-1}$$

(4) Let $r = \exp\left[\int_0^c \lambda_0(t)\exp\{\boldsymbol{Z}(t|\boldsymbol{\beta}_N)^T\theta\}dt - \int_0^c \lambda_0(t)\exp\{\boldsymbol{Z}(t|\boldsymbol{\beta}_N^{(new)})^T\theta\}dt\right]$;

(5) If $r \geqslant 1$, let $\boldsymbol{\beta}_N = \boldsymbol{\beta}_N^{(new)}$.

Otherwise let $\boldsymbol{\beta}_N = \boldsymbol{\beta}_N^{(new)}$ with probability r.

(6) Draw T distribution scale s_j from gamma $\left\{\frac{\nu+1}{2}, \frac{(Y_J - \boldsymbol{U}_j\boldsymbol{\beta}_N)^2}{2\sigma^2} + \frac{\nu}{2}\right\}$, $j = 1, \cdots, m$;

(7) Repeat steps 3 - 6 M times and save the draw of $\boldsymbol{\beta}_N$ ($M = 50$);

(8) Calculate $P(R < t|\boldsymbol{W}, \boldsymbol{\beta}_N, \boldsymbol{\theta}, \boldsymbol{\lambda}_0, R > c)$ for $c < t < c + 3$;

(9) Repeat 1,000 times steps 1 – 8.

(10) Average the values of $P(R < t|\boldsymbol{W}, \boldsymbol{\beta}_N, \boldsymbol{\theta}, \boldsymbol{\lambda}_0, R > c)$.

1.4 *Percentiles of summary statistics for warning messages*

[Table 2 about here.]

1.5 *Use of the internet calculator*

The website psacalc.sph.umich.edu contains a number of other programs as well as the one described in this paper. The specific one in this paper can be found by following the link to "Predicting future disease progression using baseline characteristics and post-treatment

PSA values, when Androgen deprivation treatment was not given." The user enters his pre-treatment PSA, Gleason and T-stage values and the date radiation therapy treatment ended. He then enters the post treatment PSA values together with their dates. Pressing "Calculate" starts the quick MCMC, and the predictions of future PSA and probability of recurrence are displayed together with any warning messages.

Received January 2008. *Revised February* 2008. *Accepted March* 2008.

Table 1: Summary statistics of posterior distributions of the population parameters.

	Variable Name	Median	Mean	SD	Lower	Upper
α_1 :	Intercept	-0.281	-0.282	0.0404	-0.362	-0.202
α_2 :	Baseline PSA	0.842	0.842	0.0179	0.807	0.878
α_3 :	Intercept	0.688	0.688	0.0912	0.51	0.865
α_4 :	I(Tstage=2)	0.360	0.360	0.0531	0.256	0.465
α_5 :	I(Tstage>2)	0.268	0.267	0.139	-0.00961	0.542
α_6 :	Baseline PSA	0.75	0.75	0.0393	0.673	0.826
α_7 :	Intercept	-0.231	-0.231	0.0299	-0.289	-0.172
α_8 :	I(Tstage=2)	0.193	0.192	0.0201	0.154	0.232
α_9 :	I(Tstage>2)	0.447	0.447	0.0541	0.338	0.553
α_{10} :	I(Gleason =7)	0.0234	0.0236	0.0177	-0.0105	0.0584
α_{11} :	I(Gleason>7)	0.135	0.135	0.0294	0.078	0.193
α_{12} :	Baseline PSA	0.178	0.178	0.0127	0.153	0.203
σ^2		0.0392	0.0392	0.000565	0.0382	0.0404
Σ_{11}		0.356	0.356	0.0146	0.328	0.385
Σ_{12}		0.267	0.267	0.0237	0.221	0.314
Σ_{13}		0.00927	0.00946	0.00684	-0.00380	0.0230
Σ_{22}		1.41	1.41	0.0635	1.29	1.54
Σ_{23}		0.275	0.276	0.0183	0.242	0.314
Σ_{33}		0.172	0.172	0.00793	0.157	0.188
θ_1	I(Tstage=2)	0.73	0.724	0.219	0.29	1.16
θ_2	I(Tstage$\geqslant$3)	0.951	0.951	0.256	0.444	1.46
θ_3	Baseline PSA	-0.269	-0.27	0.077	-0.416	-0.116
θ_4	I(Gleason =7)	0.554	0.553	0.145	0.28	0.838
θ_5	I(Gleason$\geqslant$8)	0.401	0.399	0.195	0.00887	0.768
θ_6	Expected PSA(t)	4.26	4.26	0.318	3.65	4.88
θ_7	d(Expected PSA(t))/dt	0.953	0.952	0.148	0.669	1.25
θ_8	I(t $\geqslant$ H$_i$)	-1.25	-1.25	0.183	-1.61	-0.912
$\lambda_{01} \times 10^{-3}$		1.84	1.93	0.614	0.993	3.35
$\lambda_{02} \times 10^{-3}$		1.60	1.68	0.531	0.872	2.95
$\lambda_{03} \times 10^{-3}$		1.28	1.33	0.407	0.709	2.32
$\lambda_{04} \times 10^{-3}$		1.09	1.15	0.376	0.580	2.05
$\lambda_{05} \times 10^{-3}$		1.39	1.46	0.479	0.725	2.59

Table 2: Percentiles of the statistics estimated from training data.

Statistics		95% Percentiles	99% Percentiles
$\frac{1}{m}\sum_{j=1}^{m}(Y_j - \hat{Y}_j)^2$	$m \leqslant 4$	0.234	2.293
	$m > 4$	0.204	0.534
$(Y_m - \hat{Y}_m)^2$		0.359	2.562
$\max_j (Y_j - \hat{Y}_j)^2$		1.201	4.007
$\hat{\beta}_0 - \boldsymbol{x}_0^T \hat{\boldsymbol{\alpha}}$	Upper	0.972	1.310
	Lower	-1.093	-1.730
$\hat{\beta}_1 - \boldsymbol{x}_1^T \hat{\boldsymbol{\alpha}}$	Upper	2.023	2.890
	Lower	-1.873	-2.801
$\hat{\beta}_2 - \boldsymbol{x}_2^T \hat{\boldsymbol{\alpha}}$	Upper	0.925	1.396
	Lower	-0.588	-0.853
$(\hat{\boldsymbol{\beta}} - \boldsymbol{x}^T \hat{\boldsymbol{\alpha}})^T \hat{\boldsymbol{\Sigma}}^{-1} (\hat{\boldsymbol{\beta}} - \boldsymbol{x}^T \hat{\boldsymbol{\alpha}})$		8.360	16.810
$\mathrm{var}\{E\hat{Y}(c+3)\}$		1.762	2.801