南區統計研討會

2015 / 6 / 28

A joint frailty-copula model for semicompeting risks data with meta-analysis

(in revision, *Statistical Methods In Medical Research*)

Takeshi Emura (江村剛志)

Graduate Institute of Statistics

National Central University (Taiwan)

Joint work with

Masahiro Nakatochi (Japan)

Kenta Murotani (Japan)

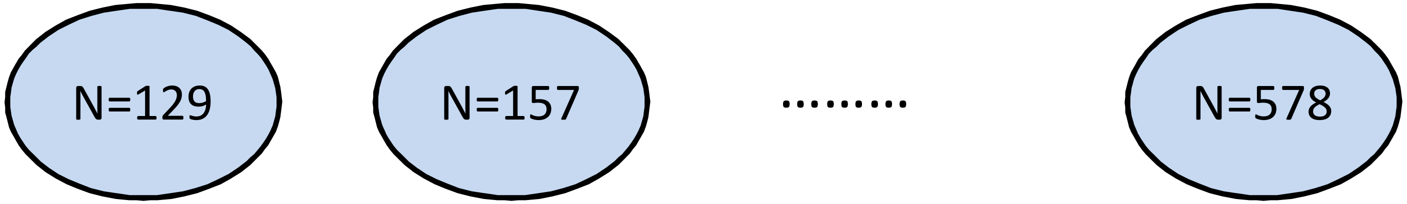
Virginie Rondeau (France)

Outline

- Review -- meta-analysis
- Review -- joint model
- Proposed method
- Simulation (skip)
- Data analysis

Meta-Analysis

- Synthesize multiple independent studies



- Useful to detect small effect size:
  - ∨ Treatment effect of chemotherapy on survival
    : Head & neck cancer (Pignon et al. 2009)
  - ∨ The effect of *CXCL12* gene on survival
    : Ovarian cancer (Ganzfried et al. 2013)
  - ∨ The effect of *ECRG4* gene on survival
    : Breast cancer (Sabatier et al. 2011)

Typical endpoints

- Time-to-progression (TTP)
  (e.g., recurrence, metastasis)
- Overall survival (OS)
  (Death from any cause)

} Event times of interest

- **Progression-free survival [PFS = min(TTP, OS)]**

Meta-analysis on event times

1) Head & neck cancer data (Pignon et al., 2000; 2009)

➔ Fit *separate* Cox models on PSF and OS, respectively

2) Ovarian cancer data (Ganzfried et al. 2013)

➔ Fit a Cox model on OS

3) Breast cancer data (Sabatier et al. 2011)

➔ Fit *separate* Cox models on PFS and OS, respectively

- Joint model = a bivariate model for TTP and OS

Time-to-progression (TTP)
Death (OS)

} Bivariate survival models (Large literature)

<u>In meta-analysis:</u>

➔ Study specific (random) effect to explain the heterogeneity

i) Bivariate survival analysis (Burzykowski et al. 2001) : (TTP, OS) is jointly observed

ii) Semi-competing risks analysis (Rondeau et al. 2011) : TTP is observed only if TTP < OS (semicompeting risks)

- Combining heterogeneous studies
  using unobserved random effect (called frailty)

$$u_i \sim f_\eta(u) = \frac{1}{\Gamma(1/\eta)\eta^{1/\eta}} u^{\frac{1}{\eta}-1} \exp\left(-\frac{u}{\eta}\right)$$

- Clustered data structure:

G independent studies ($i = 1, 2, ..., G$)

each study contain N_i subjects ($j = 1, 2, ..., N_i$)

Ref:(Burzykowski et al. 2001; Rondeau et al. 2011)

Motivation: Meta-analysis for ovarian cancer (Ganzfried et al. 2013)

A meta-analytic data of ovarian cancer patients.

Dataset[a]	Sample size	The number of observed events (event rates %)		
		Relapse ($\delta_{ij}=1$)	Death ($\delta_{ij}^{*}=1$)	Censoring ($\delta_{ij}^{*}=0$)
GSE17260	$N_1=110$	76 (69%)	46 (42%)	64 (58%)
GSE30161	$N_2=58$	48 (83%)	36 (62%)	22 (38%)
GSE9891	$N_3=278$	185 (67%)	113 (41%)	165 (59%)
TCGA	$N_4=557$	266 (48%)	290 (52%)	267 (48%)
Total	$\sum_{i=1}^{4} N_i=1003$	575 (57%)	485 (48%)	518 (52%)

- Goal of Gantzfried et al.: Survival analysis on death (OS).
- Our goal: Joint Survival analysis of relapse (TTP) and death (OS)

Data structure

X_{ij} = TTP (Time to progression due to recurrence, Relapse, etc.)

D_{ij} = OS (Overall survival = time to death from any cause)

C_{ij} = Administrative censoring time (e.g., study end)

<u>Joint frailty model</u> (Rondeau et al., 2011)

$$\begin{cases} r_{ij}(t \mid u_i) = u_i r_0(t)\exp(\boldsymbol{\beta}_1' \mathbf{Z}_{ij}) & (\text{hazard for } X_{ij}) \\ \lambda_{ij}(t \mid u_i) = u_i^{\alpha} \lambda_0(t)\exp(\boldsymbol{\beta}_2' \mathbf{Z}_{ij}) & (\text{hazard for } D_{ij}) \end{cases}$$

u_i = Study specific random effect (frailty)

$\boldsymbol{\beta}_1'$ = Effect on time-to-progression X_{ij}

$\boldsymbol{\beta}_1'$ = Effect on time-to-death D_{ij}

Data structure

X_{ij} = TTP (Time to progression due to recurrence, Relapse, etc.)

D_{ij} = OS (Overall survival = time to death from any cause)

C_{ij} = Administrative censoring time (e.g., study end)

Observations are in semicompeting risks form (Fine et al., 2001) :

$$(T_{ij} , T_{ij}^{*}, \delta_{ij} , \delta_{ij}^{*}, \mathbf{Z}_{ij}), \quad i = 1, 2, ..., G, \quad j = 1, 2, ..., N_i$$

∗ First occuring event time

$$T_{ij} = \min(X_{ij}, D_{ij}, C_{ij}), \qquad \delta_{ij} = \mathbf{I}(T_{ij} = X_{ij})$$

Indicator of progression

∗ Terminal event time

$$T_{ij}^{*} = \min(D_{ij}, C_{ij}), \qquad \delta_{ij}^{*} = \mathbf{I}(T_{ij}^{*} = D_{ij})$$

Indicator of death

Data structure

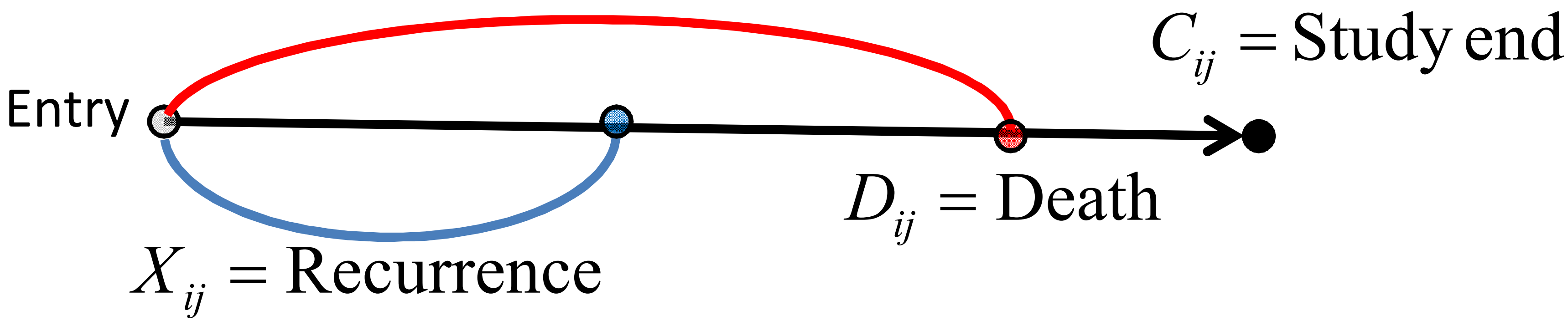


Fig. Case of $\delta_{ij} = 1, \quad \delta_{ij}^* = 1$

$*$ First occuring event time :

$$T_{ij} = \min(X_{ij}, D_{ij}, C_{ij}) = X_{ij}$$

$$\delta_{ij} = \mathbf{I}(T_{ij} = X_{ij}) = 1$$

$*$ Terminal event time :

$$T_{ij}^* = \min(D_{ij}, C_{ij}) = D_{ij}$$

$$\delta_{ij}^* = \mathbf{I}(T_{ij}^* = D_{ij}) = 1$$

4 patterns

	Likelihood contribution
• Relapse → Death T_{ij} T_{ij}^{*}	$\Pr(X_{ij}=T_{ij}, D_{ij}=T_{ij}^{*} \mid u_i)$
• Relapse → Censoring T_{ij} T_{ij}^{*}	$\Pr(X_{ij}=T_{ij}, D_{ij}>T_{ij}^{*} \mid u_i)$
• Death (without relapse) $T_{ij}=T_{ij}^{*}$	$\Pr(X_{ij}>T_{ij}, D_{ij}=T_{ij}^{*} \mid u_i)$
• Censoring (neither relapse nor death) $T_{ij}=T_{ij}^{*}$	$\Pr(X_{ij}>T_{ij}, D_{ij}>T_{ij}^{*} \mid u_i)$

log-likelihood of Rondeau et al. (2011):

$$\ell(\alpha, \eta, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, r_0, \lambda_0)$$
$$= \sum_{i=1}^{G} \left[\sum_{j=1}^{N_i} \{ \delta_{ij} \log r_{ij}(T_{ij}) + \delta_{ij}^* \log \lambda_{ij}(T_{ij}^*) \} \right.$$
$$\left. + \log \int_0^{\infty} \left\{ u_i^{m_i + \alpha m_i^*} \exp\left(-u_i \sum_{j=1}^{N_i} R_{ij}(T_{ij}) - u_i^{\alpha} \sum_{j=1}^{N_i} \Lambda_{ij}(T_{ij}^*) \right) \right\} f_{\eta}(u_i) du_i \right],$$

where $m_i = \sum_{j=1}^{N_i} \delta_{ij}$ and $m_i^* = \sum_{j=1}^{N_i} \delta_{ij}^*$.

- Nonparametric hazard approximation via **Cubic M-Spline**

(O' Sullivan 1988; Joly, Commenges and Letenneur 1998)

$$r_0(t) = \sum_{\ell=1}^{L_r} g_{\ell} M_{\ell}(t), \qquad \lambda_0(t) = \sum_{\ell=1}^{L_{\lambda}} h_{\ell} M_{\ell}(t)$$

Proposed Idea

$$X_{ij} = \text{TTP} \ (\text{Recurrence, Relapse, etc.})$$

$$D_{ij} = \text{OS} \ (\text{Death from any cause})$$

Joint frailty model (Rondeau et al., 2011)

$$\begin{cases} r_{ij}(t \mid u_i) = u_i r_0(t) \exp(\boldsymbol{\beta}_1' \mathbf{Z}_{ij}) & (\text{time-to-progression } X_{ij}) \\ \lambda_{ij}(t \mid u_i) = u_i^{\alpha} \lambda_0(t) \exp(\boldsymbol{\beta}_2' \mathbf{Z}_{ij}) & (\text{time-to-death } D_{ij}) \end{cases}$$

- They assume Independence within a subject *j* :

$$X_{ij} \perp D_{ij} \mid u_i$$

➔ **Our proposed idea:**
Relax this intra-subject independence assumption via Copulas

Joint frailty-copula model (Proposed)

Joint frailty model (Rondeau et al., 2011)

$$\begin{cases} r_{ij}(t \mid u_i) = u_i r_0(t)\exp(\boldsymbol{\beta}_1'\mathbf{Z}_{ij}) & (\text{ time-to-progression } X_{ij}) \\ \lambda_{ij}(t \mid u_i) = u_i^{\alpha}\lambda_0(t)\exp(\boldsymbol{\beta}_2'\mathbf{Z}_{ij}) & (\text{ time-to-death } D_{ij}) \end{cases}$$

+

Copula model (new):

$$\Pr(X_{ij} > x, D_{ij} > y \mid u_i) = C_\theta[\exp\{-R_{ij}(x \mid u_i)\}, \exp\{-\Lambda_{ij}(y \mid u_i)\}]$$

C_θ : Copula (Nelsen, 2006), $R_{ij}(x \mid u_i) = \int_0^x r_{ij}(v \mid u_i)dv \quad \Lambda_{ij}(y \mid u_i) = \int_0^y \lambda_{ij}(v \mid u_i)dv$

Copula parameter θ

➔ Intra-subject Kendall's tau $\tau_\theta(X_{ij}, D_{ij} \mid u_i)$

e.g., Clayton copula: $\tau(X_{ij}, D_{ij} \mid u_i) = \theta/(\theta + 2)$

Clustered semicompeting risks data (same as Rondeau et al., 2011)

$(T_{ij} , T_{ij}^{*} , \delta_{ij} , \delta_{ij}^{*})$ for

$i = 1, 2, ..., G$ and $j = 1, 2, ..., N_i$.

First occurring event	T_{ij}	T_{ij}^{*}	δ_{ij}	δ_{ij}^{*}	Likelihood contribution
Progression	X_{ij}	D_{ij}	1	1	$\Pr(X_{ij} = T_{ij}, D_{ij} = T_{ij}^{*} \mid u_i)$
Progression	X_{ij}	C_{ij}	1	0	$\Pr(X_{ij} = T_{ij}, D_{ij} > T_{ij}^{*} \mid u_i)$
Death	D_{ij}	D_{ij}	0	1	$\Pr(X_{ij} > T_{ij}, D_{ij} = T_{ij}^{*} \mid u_i)$
Censoring	C_{ij}	C_{ij}	0	0	$\Pr(X_{ij} > T_{ij}, D_{ij} > T_{ij}^{*} \mid u_i)$

Log-likelihood (proposed)

$$\ell(\alpha,\eta,\theta,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,r_0,\lambda_0)$$

$$= \sum_{i=1}^{G}\Bigg[\sum_{j=1}^{N_i}\{\delta_{ij}\log r_{ij}(T_{ij})+\delta_{ij}^*\log\lambda_{ij}(T_{ij}^*)\}$$

$$+\log\int_0^\infty\Bigg\{u_i^{m_i+\alpha m_i^*}\prod_{j=1}^{N_i}\psi_\theta[u_iR_{ij}(T_{ij}),u^\alpha\Lambda_{ij}(T_{ij}^*)]^{\delta_{ij}}\psi_\theta^*[u_iR_{ij}(T_{ij}),u_i^\alpha\Lambda_{ij}(T_{ij}^*)]^{\delta_{ij}^*}$$

$$\times\Theta_\theta[u_iR_{ij}(T_{ij}),u_i^\alpha\Lambda_{ij}(T_{ij}^*)]^{\delta_{ij}\delta_{ij}^*}D_\theta[u_iR_{ij}(T_{ij}),u_i^\alpha\Lambda_{ij}(T_{ij}^*)]\Bigg\}f_\eta(u_i)du_i\Bigg],$$

where $r_{ij}(t)=r_0(t)\exp(\boldsymbol{\beta}_1'\mathbf{Z}_{ij})$, $\lambda_{ij}(t)=\lambda_0(t)\exp(\boldsymbol{\beta}_2'\mathbf{Z}_{ij})$,

$r_0(t)=\sum_{\ell=1}^{L_r}g_\ell M_\ell(t)$, $\lambda_0(t)=\sum_{\ell=1}^{L_\lambda}h_\ell M_\ell(t)$, ⇦ Cubic M-spline

$D_\theta[s,t]=C_\theta[\exp(-s),\exp(-t)]$, $\psi_\theta=D_\theta^{[1,0]}/D_\theta$, $D_\theta^{[1,0]}=-\partial D_\theta/\partial s$, $\psi_\theta^*=D_\theta^{[0,1]}/D_\theta$,

$D_\theta^{[0,1]}=-\partial D_\theta/\partial t$, $\Theta_\theta=D_\theta^{[1,1]}D_\theta/D_\theta^{[1,0]}D_\theta^{[0,1]}$ and $D_\theta^{[1,1]}=\partial^2 D_\theta/\partial s\partial t$.

⇧ Derivatives of copula

Log-likelihood (proposed)

- Independent copula $C_\theta(v, w) = vw$

➔ Reduces to the log-likelihood of Rondeau et al. (2011):

$$\ell(\alpha, \eta, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, r_0, \lambda_0) = \sum_{i=1}^{G} \left[\sum_{j=1}^{N_i} \{ \delta_{ij} \log r_{ij}(T_{ij}) + \delta_{ij}^* \log \lambda_{ij}(T_{ij}^*) \} \right.$$

$$\left. + \log \int_0^\infty \left\{ u_i^{m_i + \alpha m_i^*} \exp\left(-u_i \sum_{j=1}^{N_i} R_{ij}(T_{ij}) - u_i^\alpha \sum_{j=1}^{N_i} \Lambda_{ij}(T_{ij}^*) \right) \right\} f_\eta(u_i) du_i \right].$$

where $m_i = \sum_{j=1}^{N_i} \delta_{ij}$ and $m_i^* = \sum_{j=1}^{N_i} \delta_{ij}^*$.

- Penalized likelihood with cubic M-spline

➔ Directly follow Rondeau et al. (2011)

$$\ell(\,\alpha,\eta,\theta,\mathbf{\beta}_1,\mathbf{\beta}_2,r_0,\lambda_0\,)-\kappa_1\int\ddot{\gamma}_0(t)^2\,dt-\kappa_2\int\ddot{\lambda}_0(t)^2\,dt$$

$$\int\ddot{r}_0(t)^2\,dt=\sum_{k=1}^{L_r}\sum_{\ell=1}^{L_r}g_k g_\ell\int\ddot{M}_k(t)\ddot{M}_\ell(t)dt,\quad \int\ddot{\lambda}_0(t)^2\,dt=\sum_{k=1}^{L_\lambda}\sum_{\ell=1}^{L_\lambda}h_k h_\ell\int\ddot{M}_k(t)\ddot{M}_\ell(t)dt$$

- κ_1 = Smoothing parameter for the hazard of TTP
- κ_2 = Smoothing parameter for the hazard of OS

The values $(\,\hat{\kappa}_1,\hat{\kappa}_2\,)$ chosen by LCV (Joly, et al. 1998)

$$(\,\hat{\alpha},\hat{\eta},\hat{\theta},\hat{\mathbf{\beta}}_1,\hat{\mathbf{\beta}}_2,\hat{r}_0,\hat{\lambda}_0\,)$$
$$=\underset{(\alpha,\eta,\theta,\mathbf{\beta}_1,\mathbf{\beta}_2,r_0,\lambda_0)}{\arg\max}\left[\ell(\,\alpha,\eta,\theta,\mathbf{\beta}_1,\mathbf{\beta}_2,r_0,\lambda_0\,)-\hat{\kappa}_1\int\ddot{\gamma}_0(t)^2\,dt-\hat{\kappa}_2\int\ddot{\lambda}_0(t)^2\,dt\right]$$

•Automatic computing implemented in our R *joint.Cox* package
Accuracy of the package checked by simulations (in our paper)

Ovarian cancer meta-analysis

(Ganzfried et al. 2013)

Dataset[a]	Sample size	The number of observed events (event rates %)		
		Relapse ($\delta_{ij}=1$)	Death ($\delta_{ij}^{*}=1$)	Censoring ($\delta_{ij}^{*}=0$)
GSE17260	$N_1=110$	76 (69%)	46 (42%)	64 (58%)
GSE30161	$N_2=58$	48 (83%)	36 (62%)	22 (38%)
GSE9891	$N_3=278$	185 (67%)	113 (41%)	165 (59%)
TCGA	$N_4=557$	266 (48%)	290 (52%)	267 (48%)
Total	$\sum_{i=1}^{4} N_i=1003$	575 (57%)	485 (48%)	518 (52%)

- Goal 1: Marginal analysis of relapse (TTP) and death (OS)

$$\begin{cases} r_{ij}(t\,|\,u_i)=u_i r_0(t)\exp(\ \beta_1\times\text{CXCL12}\ ) & (\text{ hazard for TTP }) \\ \lambda_{ij}(t\,|\,u_i)=u_i^{\alpha}\lambda_0(t)\exp(\ \beta_2\times\text{CXCL12}\ ) & (\text{ hazard for OS }) \end{cases}$$

- Goal 2: Association analysis of TTP and death

Copula model:

$$\Pr(\,X_{ij}>x\,,\,D_{ij}>y\,|\,u_i\,)=C_{\theta}[\,\exp\{-R_{ij}(x\,|\,u_i)\},\exp\{-\Lambda_{ij}(y\,|\,u_i)\}\,]$$

Table 5. The joint analysis of recurrence (TTP) and death (OS) for the meta-analytic data (four studies, 1003 patients) for ovarian cancer patients of Ganzfried et al.[19].

	Proposed method: Estimate (95% CI)	Method of Rondeau et al.[6]: Estimate (95% CI)
RR[a] for relapse (TTP) : $\exp(\beta_1)$	1.22 (1.13-1.32)	1.24 (1.14-1.35)
RR[a] for death (OS) : $\exp(\beta_2)$	1.18 (1.08-1.29)	1.17 (1.07-1.29)
Heterogeneity: $\eta = Var_\eta(u_i)$	0.033 (0.006-0.186)	0.028 (0.004-0.180)
Copula parameter: θ	2.35 (1.90-2.90)	0.00 (assumed fixed)
RR for death after relapse: $\theta + 1$	3.35 (2.90-3.90)	1.00 (assumed fixed)
Kendall's tau: $\tau = \theta/(\theta + 2)$	0.54 (0.49-0.59)	0.00 (assumed fixed)
Maximum penalized log-likelihood	-8604.093	-8744.023

Notes: [a]RR (Relative Risk) of *CXCL12* expression on the hazards are examined.

* Ganzfried et al. (2013) reported RR=1.15 (1.09-1.23) for OS based on 14 studies

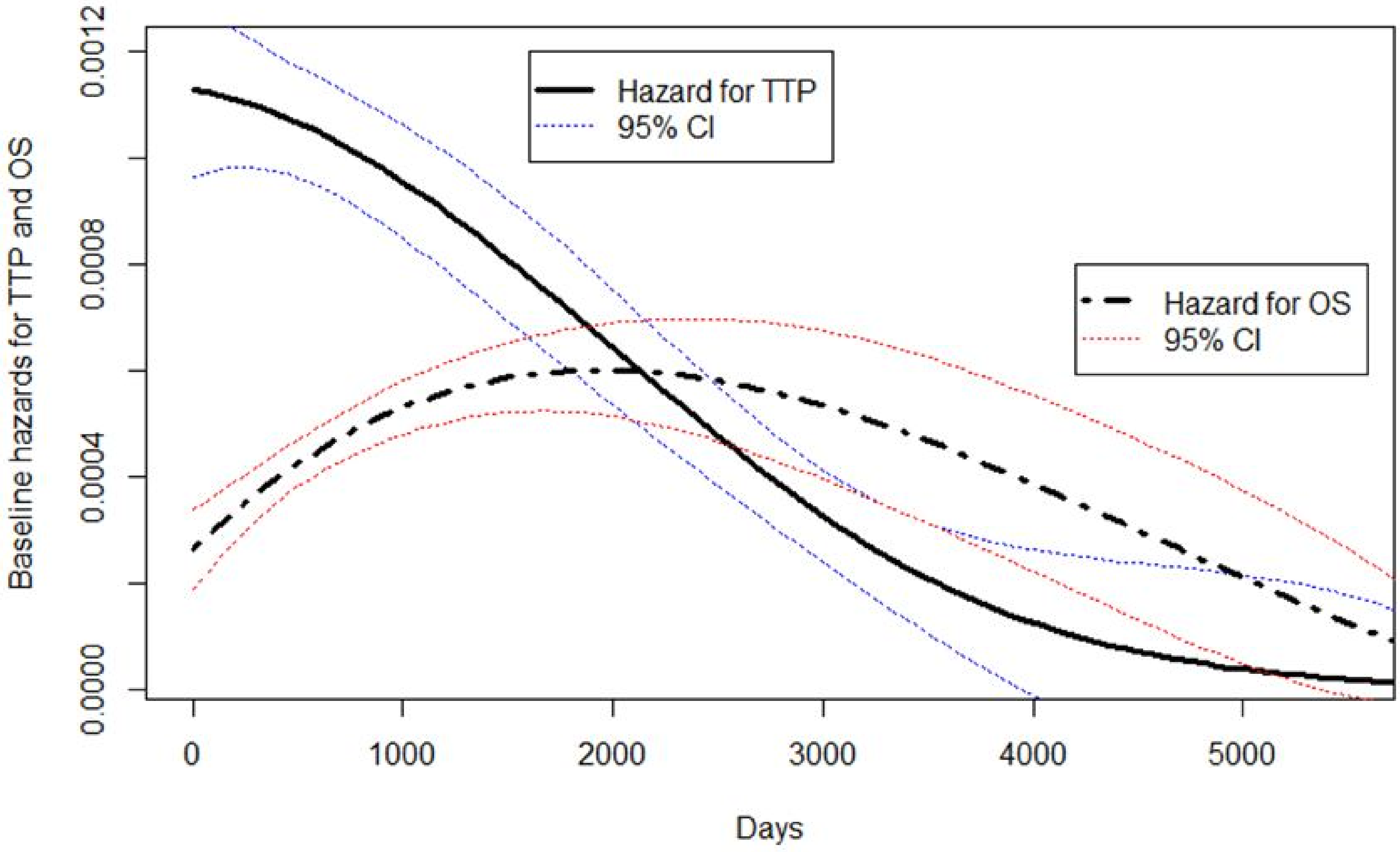


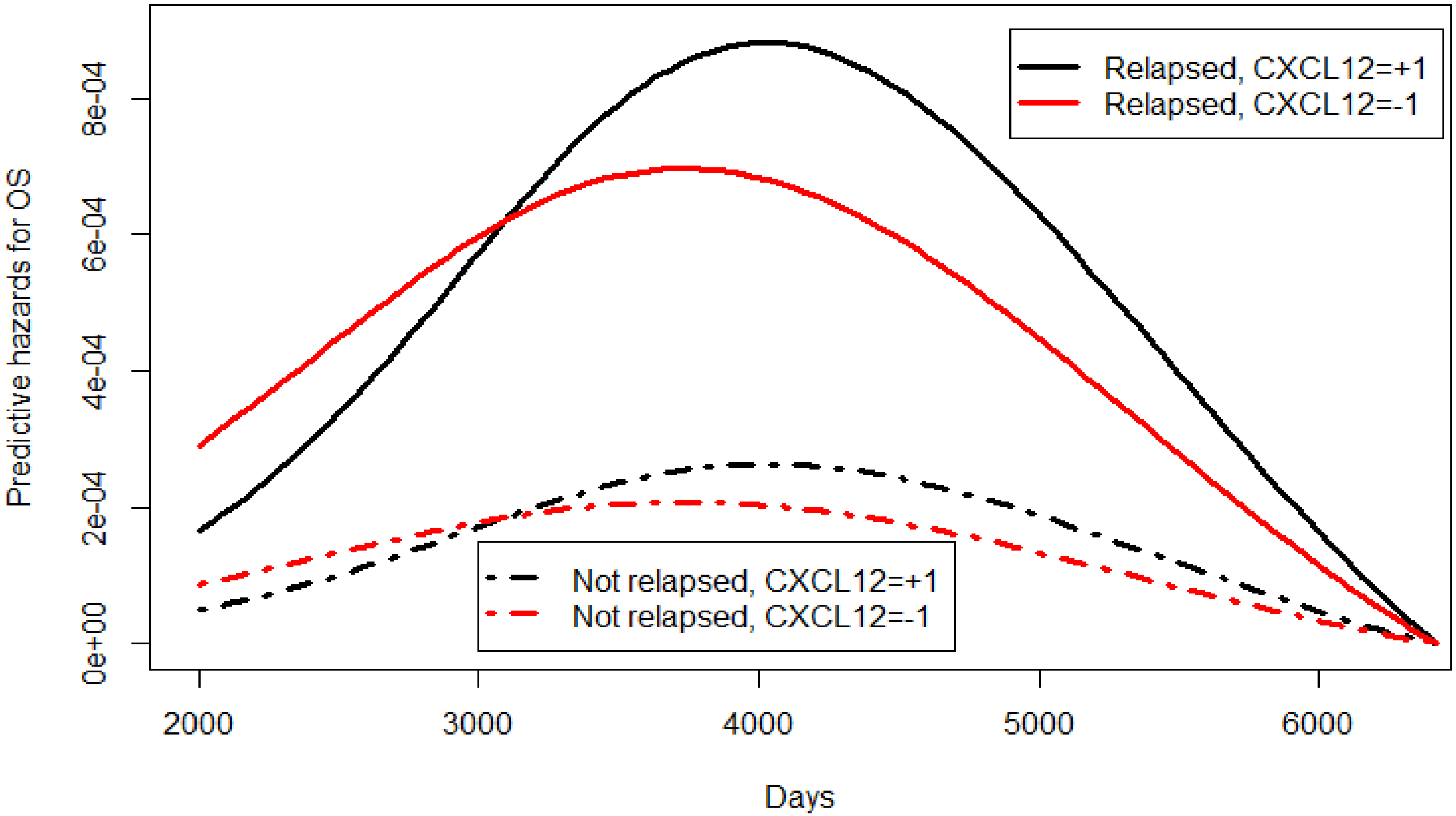
Figure 2. Baseline hazard functions for TTP (recurrence) and OS (death) based on the meta-analytic data of ovarian cancer patients. The dotted lines (red or blue color) show the 95% confidence intervals.

Copula parameter: θ =2.35 (Kendall's tau=0.54)

<u>Practical implication:</u>

Relapse occurring before death can increase the risk of death by 3.35 times:

$$3.35 \approx \theta + 1 = \frac{\lambda_{ij}(\, y \mid X_{ij} = x, Z_{ij}, u_i\,)}{\lambda_{ij}(\, y \mid X_{ij} > x, Z_{ij}, u_i\,)}, \qquad y \geq x, \qquad Z_{ij} = \text{CXCL12}$$



Summary

Propose a **joint** frailty-copula model for

Dependence between **TTP** and **OS**

- Extend the joint frailty model of Rondeau et al. (2011)

  → allow intra-subject dependence via copulas

- Data analysis:

  Shows the significant intra-subject dependence

  Risk prediction of death (OS) by prior relapse (TTP)

Future work : Dynamic prediction of death (OS) by prior relapse (TTP) as in Mauguen et al. (2013, 2015)

Visit of Dr. Virginie Rondeau in Taiwan

- Seminar talk on 7/22(Wed) 10-11AM @Academia Sinica

Thank you !