Proportional Rates Models for Multivariate Panel Count Data

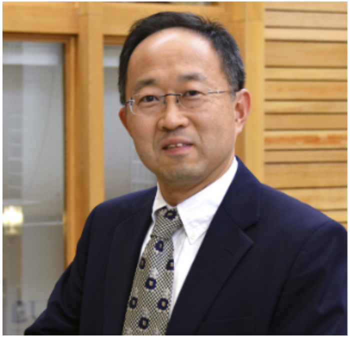
Yangjianchen Xu

Department of Biostatistics
The University of North Carolina at Chapel Hill

August 8, 2023

Acknowledgements

Dr. Donglin Zeng



Dr. Danyu Lin

Outline

Outline

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

Multivariate Panel Count Data

- recurrent events examined periodically (interval censoring)
- multiple types of events (not competing risk)

Introduction

THE UNIVERSITY *of* NORTH CAROLINA *at* CHAPEL HILL

Multivariate Panel Count Data

- recurrent events examined periodically (interval censoring)
- multiple types of events (not competing risk)

Examples

- the number of clinically and radiologically damaged joints in a psoriatic arthritis patient (Gladman et al., 1995)
- the number of basal and squamous cell tumors in a skin cancer patient (Bailey et al., 2010)

Introduction

Multivariate Panel Count Data

- recurrent events examined periodically (interval censoring)
- multiple types of events (not competing risk)

Examples

- the number of clinically and radiologically damaged joints in a psoriatic arthritis patient (Gladman et al., 1995)
- the number of basal and squamous cell tumors in a skin cancer patient (Bailey et al., 2010)

Theoretical/Computational Issues

- no exact failure time
- complex dependence for the recurrent events of the same type and of different types

Existing methods

Random-effects models

- Zeng and Lin (2020) and the references therein

Existing methods

Random-effects models

- Zeng and Lin (2020) and the references therein

Marginal models: proportional rates/means models

- Sun and Wei (2000), He et al. (2007): independent or modeled examination times
- Wellner and Zhang (2007): slow and unstable doubly iterative algorithm
- Lu et al. (2009): arbitrary choices of spline functions

Our contributions

THE UNIVERSITY *of* NORTH CAROLINA *at* CHAPEL HILL

- a simple and stable EM-type algorithm is used for estimation

Our contributions

- a simple and stable EM-type algorithm is used for estimation
- examination times are not modeled

Our contributions

- a simple and stable EM-type algorithm is used for estimation
- examination times are not modeled
- cumulative baseline rate functions are estimated nonparametrically

Our contributions

- a simple and stable EM-type algorithm is used for estimation
- examination times are not modeled
- cumulative baseline rate functions are estimated nonparametrically
- asymptotic theory is established

Our contributions

- a simple and stable EM-type algorithm is used for estimation
- examination times are not modeled
- cumulative baseline rate functions are estimated nonparametrically
- asymptotic theory is established
- graphical and numerical model checking techniques are first proposed

Outline

Notation

- n = number of subjects
- K = number of types of events
- $X_i(\cdot)$ = (potentially time-dependent) covariates
- $N_{ki}(\cdot)$ = counting process of the kth type of event for the ith subject
- $0 < U_{ki1} < \cdots < U_{ki,m_{ki}} = C_{ki}$ are examination times for $N_{ki}(\cdot)$
- $\Delta_{kij} = N_{ki}(U_{kij}) - N_{ki}(U_{ki,j-1})$ $(j = 1, \ldots, m_{ki})$

Models

Proportional Rates Models

$$E\{\mathrm{d}N_{ki}(t) \mid X_i(t)\} = \exp\{\beta_k^{\mathrm{T}} X_i(t)\}\mathrm{d}\Lambda_k(t)$$

- $\mathrm{d}N_{ki}(t) = N_{ki}\{(t + \mathrm{d}t)-\} - N_{ki}(t-)$
- β_k = regression parameters
- $\Lambda_k(t)$ = arbitrary non-decreasing baseline cumulative rate function

Models

Proportional Rates Models

$$E\{\mathrm{d}N_{ki}(t) \mid X_i(t)\} = \exp\{\beta_k^{\mathrm{T}} X_i(t)\}\mathrm{d}\Lambda_k(t)$$

- $\mathrm{d}N_{ki}(t) = N_{ki}\{(t + \mathrm{d}t)-\} - N_{ki}(t-)$
- β_k = regression parameters
- $\Lambda_k(t)$ = arbitrary non-decreasing baseline cumulative rate function

Working Assumptions

- all types of event times are independent
- $N_{ki}(t)$ is a nonhomogeneous Poisson process
  - Δ_{kij} are independent Poisson with means $\int_{U_{ki,j-1}}^{U_{kij}} \exp\left\{\beta_k^{\mathrm{T}} X_i(u)\right\} \mathrm{d}\Lambda_k(u)$

Estimation procedure

Pseudo-Likelihood

$$\prod_{i=1}^{n}\left(\prod_{j=1}^{m_{ki}}\frac{\left[\int_{U_{ki,j-1}}^{U_{kij}}\exp\{\beta_k^{\mathrm{T}}X_i(t)\}\mathrm{d}\Lambda_k(t)\right]^{\Delta_{kij}}}{\Delta_{kij}!}\right)\exp\left[-\int_0^{C_{ki}}\exp\{\beta_k^{\mathrm{T}}X_i(t)\}\mathrm{d}\Lambda_k(t)\right]$$

Estimation procedure

Pseudo-Likelihood

$$\prod_{i=1}^{n}\left(\prod_{j=1}^{m_{ki}}\frac{\left[\int_{U_{ki,j-1}}^{U_{kij}}\exp\{\beta_k^{\mathrm{T}}X_i(t)\}\mathrm{d}\Lambda_k(t)\right]^{\Delta_{kij}}}{\Delta_{kij}!}\right)\exp\left[-\int_0^{C_{ki}}\exp\{\beta_k^{\mathrm{T}}X_i(t)\}\mathrm{d}\Lambda_k(t)\right]$$

Nonparametric Maximum Pseudo-Likelihood Estimation

- $t_{k1} < \cdots < t_{kd_k}$ = the unique values of all examination times
- λ_{kl} = jump size of $\Lambda_k(\cdot)$ at t_{kl}
- For each k, we maximize

$$\prod_{i=1}^{n}\left[\prod_{j=1}^{m_{ki}}\frac{\left\{\sum_{l:t_{kl}\in(U_{ki,j-1},U_{kij}]}\lambda_{kl}\exp(\beta_k^{\mathrm{T}}X_{kil})\right\}^{\Delta_{kij}}}{\Delta_{kij}!}\right]\exp\left\{-\sum_{l:t_{kl}\leqslant C_{ki}}\lambda_{kl}\exp(\beta_k^{\mathrm{T}}X_{kil})\right\},$$

where $X_{kil} = X_i(t_{kl})$.

EM-type algorithm

Missing data (latent variables): $W_{kil} \overset{\text{ind}}{\sim} \text{Poisson}\left(\lambda_{kl} e^{\beta_k^{\mathrm{T}} X_{kil}}\right)$

EM-type algorithm

Missing data (latent variables): $W_{kil} \overset{\text{ind}}{\sim} \text{Poisson}\left(\lambda_{kl} e^{\beta_k^{\mathrm{T}} X_{kil}}\right)$

Observed data: $\{X_i(\cdot), \sum_{l: t_{kl} \in (U_{ki,j-1}, U_{kij}]} W_{kil} = \Delta_{kij}\}$

EM-type algorithm

Missing data (latent variables): $W_{kil} \overset{\text{ind}}{\sim} \text{Poisson}\left(\lambda_{kl} e^{\beta_k^{\mathrm{T}} X_{kil}}\right)$

Observed data: $\{X_i(\cdot), \sum_{l: t_{kl} \in (U_{ki,j-1}, U_{kij}]} W_{kil} = \Delta_{kij}\}$

Observed-data likelihood

$$\prod_{i=1}^{n} \prod_{j=1}^{m_{ki}} \Pr\left(\sum_{l: t_{kl} \in (U_{ki,j-1}, U_{kij}]} W_{kil} = \Delta_{kij} \right) = \textbf{Pseudo-Likelihood}$$

EM-type algorithm

Missing data (latent variables): $W_{kil} \overset{\text{ind}}{\sim} \text{Poisson}\left(\lambda_{kl} e^{\beta_k^{\mathrm{T}} X_{kil}}\right)$

Observed data: $\{X_i(\cdot), \sum_{l:t_{kl}\in(U_{ki,j-1},U_{kij}]} W_{kil} = \Delta_{kij}\}$

Observed-data likelihood

$$\prod_{i=1}^{n}\prod_{j=1}^{m_{ki}} \Pr\left(\sum_{l:t_{kl}\in(U_{ki,j-1},U_{kij}]} W_{kil} = \Delta_{kij}\right) = \textbf{Pseudo-Likelihood}$$

Complete-data log-likelihood

$$\sum_{i=1}^{n}\sum_{l=1}^{d_k} I\left(t_{kl} \leqslant C_{ki}\right)\left\{W_{kil}(\log\lambda_{kl} + \beta_k^{\mathrm{T}} X_{kil}) - \lambda_{kl}\exp(\beta_k^{\mathrm{T}} X_{kil}) - \log W_{kil}!\right\}$$

EM-type algorithm

E-step

$$\widehat{E}\left(W_{kil}\right) = I(U_{ki,j-1} < t_{kl} \leqslant U_{kij}) \frac{\Delta_{kij} \lambda_{kl} \exp(\beta_k^{\mathrm{T}} X_{kil})}{\sum_{s: t_{ks} \in (U_{ki,j-1}, U_{kij}]} \lambda_{ks} \exp(\beta_k^{\mathrm{T}} X_{kis})}$$

EM-type algorithm

E-step

$$\widehat{E}\left(W_{kil}\right) = I(U_{ki,j-1} < t_{kl} \leqslant U_{kij}) \frac{\Delta_{kij}\lambda_{kl}\exp(\beta_k^{\mathrm{T}} X_{kil})}{\sum_{s:t_{ks}\in(U_{ki,j-1},U_{kij}]} \lambda_{ks}\exp(\beta_k^{\mathrm{T}} X_{kis})}$$

M-step

$$\sum_{i=1}^{n}\sum_{l=1}^{d_k} I(C_{ki} \geqslant t_{kl})\widehat{E}(W_{kil})\left\{X_{kil} - \frac{\sum_{i'=1}^{n} I(C_{ki'} \geqslant t_{kl})\exp(\beta_k^{\mathrm{T}} X_{ki'l})X_{ki'l}}{\sum_{i'=1}^{n} I(C_{ki'} \geqslant t_{kl})\exp(\beta_k^{\mathrm{T}} X_{ki'l})}\right\} = 0.$$

We then update

$$\lambda_{kl} = \frac{\sum_{i=1}^{n} I(C_{ki} \geqslant t_{kl})\widehat{E}\left(W_{kil}\right)}{\sum_{i=1}^{n} I(C_{ki} \geqslant t_{kl})\exp\left(\beta_k^{\mathrm{T}} X_{kil}\right)}$$

Asymptotic properties

Write $\widehat{\beta} = (\widehat{\beta}_1^{\mathrm{T}}, \ldots, \widehat{\beta}_K^{\mathrm{T}})^{\mathrm{T}}$ and $\widehat{\Lambda} = (\widehat{\Lambda}_1, \ldots, \widehat{\Lambda}_K)$.

Asymptotic properties

Write $\widehat{\beta} = (\widehat{\beta}_1^{\mathrm{T}}, \ldots, \widehat{\beta}_K^{\mathrm{T}})^{\mathrm{T}}$ and $\widehat{\Lambda} = (\widehat{\Lambda}_1, \ldots, \widehat{\Lambda}_K)$.

Consistency

Theorem 1

Under some regularity conditions, $\|\widehat{\beta} - \beta_0\| + \sum_{k=1}^{K} \sup_{t \in [0, \tau_k]} |\widehat{\Lambda}_k(t) - \Lambda_{0k}(t)| \to 0$ *almost surely, where* $\| \cdot \|$ *is the Euclidean norm.*

Asymptotic distribution

Theorem 2

Under some regularity conditions, $n^{1/2}(\widehat{\beta} - \beta_0)$ *converges in distribution to a zero-mean multivariate normal random vector with covariance matrix*

$$\Omega = \Sigma(\beta_0, \Lambda_0)^{-1} E\{S(\beta_0, \Lambda_0) S(\beta_0, \Lambda_0)^{\mathrm{T}}\} \Sigma(\beta_0, \Lambda_0)^{-1}.$$

Sandwich variance estimator

Profile pseudo-log-likelihood for β_k

$$\mathrm{pl}_k(\beta_k) = \sum_{i=1}^{n} \sum_{j=1}^{m_{ki}} \left[\Delta_{kij} \log \left\{ \sum_{U_{ki,j-1} < t_{kl} \leqslant U_{kij}} \widetilde{\lambda}_{kl} \exp(\beta_k^{\mathrm{T}} X_{kil}) \right\} - \sum_{U_{ki,j-1} < t_{kl} \leqslant U_{kij}} \widetilde{\lambda}_{kl} \exp(\beta_k^{\mathrm{T}} X_{kil}) \right]$$

- $\widetilde{\lambda}_{kl}$ $(l = 1, \ldots, d_k)$ are obtained from EM with fixed β_k

Sandwich variance estimator

Profile pseudo-log-likelihood for β_k

$$\mathrm{pl}_k(\beta_k) = \sum_{i=1}^{n} \sum_{j=1}^{m_{ki}} \left[\Delta_{kij} \log \left\{ \sum_{U_{ki,j-1} < t_{kl} \leqslant U_{kij}} \widetilde{\lambda}_{kl} \exp(\beta_k^{\mathrm{T}} X_{kil}) \right\} - \sum_{U_{ki,j-1} < t_{kl} \leqslant U_{kij}} \widetilde{\lambda}_{kl} \exp(\beta_k^{\mathrm{T}} X_{kil}) \right]$$

- $\widetilde{\lambda}_{kl}$ $(l = 1, \ldots, d_k)$ are obtained from EM with fixed β_k

Covariance matrix estimator between $\widehat{\beta}_k$ and $\widehat{\beta}_l$

$$\widehat{V}_{kl} = \left\{ D_{h_n}^2 \mathrm{pl}_k(\widehat{\beta}_k) \right\}^{-1} \sum_{i=1}^{n} D_{h_n} \mathrm{pl}_{ki}(\widehat{\beta}_k) D_{h_n} \mathrm{pl}_{li}(\widehat{\beta}_l)^{\mathrm{T}} \left\{ D_{h_n}^2 \mathrm{pl}_l(\widehat{\beta}_l) \right\}^{-1}$$

- $\mathrm{pl}_{ki}(\beta_k)$ = contribution of the ith subject to $pl_k(\beta_k)$

Sandwich variance estimator

Theorem 3

Under some regularity conditions, $\{n(\widehat{V}_{kl}); 1 \leqslant k, l \leqslant K\}$ is a consistent estimator for the limiting covariance matrix Ω.

Sandwich variance estimator

Theorem 3

Under some regularity conditions, $\{n(\widehat{V}_{kl}); 1 \leqslant k, l \leqslant K\}$ *is a consistent estimator for the limiting covariance matrix* Ω.

Statistical Inference

$$L\widehat{\beta} \sim N\left(L\beta, LVL'\right)$$

$$V = \begin{bmatrix} V_{11} & \cdots & V_{1K} \\ \vdots & \vdots & \vdots \\ V_{K1} & \cdots & V_{KK} \end{bmatrix}$$

- linear combinations (e.g., a subset of parameters, difference of two parameters)

Model checking procedures

Model checking procedures

Residual process

$$\widehat{M}_i(t) = N_i(t \wedge C_i) - \int_0^{t \wedge C_i} \exp(\widehat{\beta}^{\mathrm{T}} X_i) \mathrm{d}\widehat{\Lambda}(u)$$

- observed − predicted

Model checking procedures

Residual process

$$\widehat{M}_i(t) = N_i(t \wedge C_i) - \int_0^{t \wedge C_i} \exp(\widehat{\beta}^{\mathrm{T}} X_i)\mathrm{d}\widehat{\Lambda}(u)$$

- observed − predicted
- not fully observed
- $\widehat{\Lambda}$ is only $n^{1/3}$-consistent

Model checking procedures

Residual process

$$\widehat{M}_i(t) = N_i(t \wedge C_i) - \int_0^{t \wedge C_i} \exp(\widehat{\beta}^{\mathrm{T}} X_i) \mathrm{d}\widehat{\Lambda}(u)$$

- observed − predicted
- not fully observed
- $\widehat{\Lambda}$ is only $n^{1/3}$-consistent

Counting process of the examination times

$$\widetilde{N}_i(t) = \sum_{j=1}^{m_i} I(U_{ij} \leqslant t)$$

- model its rate function by $E\{\mathrm{d}\widetilde{N}_i(t) \mid X_i\} = \exp(\gamma^{\mathrm{T}} X_i)\theta(t)\mathrm{d}t$
- $n^{1/2}$-consistent estimator $\widehat{\gamma}$ from Lin et al. (2000)

Model checking procedures

Multi-parameter process

$$W(t,x) = n^{-1/2} \sum_{i=1}^{n} \int_0^t \left\{ g(u,x,X_i,\widehat{\beta}) - \widehat{h}(u,x) \right\} \widehat{M}_i(u) \mathrm{d}\widetilde{N}_i(u)$$

Model checking procedures

Multi-parameter process

$$W(t,x) = n^{-1/2}\sum_{i=1}^{n}\int_0^t \left\{g(u,x,X_i,\widehat{\beta}) - \widehat{h}(u,x)\right\} \widehat{M}_i(u)\mathrm{d}\widetilde{N}_i(u)$$

- g is chosen to check different aspects of the model
- a measure of the correlation between g and $\widehat{M}_i(\cdot)$

Model checking procedures

Multi-parameter process

$$W(t,x) = n^{-1/2} \sum_{i=1}^{n} \int_0^t \left\{ g(u,x,X_i,\widehat{\beta}) - \widehat{h}(u,x) \right\} \widehat{M}_i(u) \mathrm{d}\widetilde{N}_i(u)$$

- g is chosen to check different aspects of the model
- a measure of the correlation between g and $\widehat{M}_i(\cdot)$
- $\widetilde{N}_i(\cdot)$ is introduced to guarantee that only the observed values of $\widehat{M}_i(\cdot)$ are used
- $\widehat{h} = \frac{\sum_{j=1}^{n} g(u,x,X_j,\widehat{\beta}) \exp\{(\widehat{\beta}+\widehat{\gamma})^{\mathrm{T}} X_j\}}{\sum_{j=1}^{n} \exp\{(\widehat{\beta}+\widehat{\gamma})^{\mathrm{T}} X_j\}}$ eliminates the effects of the slow convergence of $\widehat{\Lambda}$

Model checking procedures

Multi-parameter process

$$W(t,x) = n^{-1/2} \sum_{i=1}^{n} \int_0^t \left\{ g(u,x,X_i,\widehat{\beta}) - \widehat{h}(u,x) \right\} \widehat{M}_i(u) \mathrm{d}\widetilde{N}_i(u)$$

- g is chosen to check different aspects of the model
- a measure of the correlation between g and $\widehat{M}_i(\cdot)$
- $\widetilde{N}_i(\cdot)$ is introduced to guarantee that only the observed values of $\widehat{M}_i(\cdot)$ are used
- $\widehat{h} = \frac{\sum_{j=1}^{n} g(u,x,X_j,\widehat{\beta}) \exp\{(\widehat{\beta}+\widehat{\gamma})^{\mathrm{T}} X_j\}}{\sum_{j=1}^{n} \exp\{(\widehat{\beta}+\widehat{\gamma})^{\mathrm{T}} X_j\}}$ eliminates the effects of the slow convergence of $\widehat{\Lambda}$

Zero-mean Gaussian process

$$\widetilde{W}(t,x) = n^{-1/2} \sum_{i=1}^{n} A_i(t,x,\widehat{\Lambda},\widehat{\beta},\widehat{\gamma}) G_i,$$

where G_i are independent standard normal random variables.

Model checking procedures

Choices of $g(u, x, X_i, \widehat{\beta})$

- functional form: $I(X_{iq} \leqslant x)$
- proportional means assumption: X_{iq}
- exponential link function: $I(\widehat{\beta}^{\mathrm{T}} X_i \leqslant x)$
- overall fit: $I(X_i \leqslant x)$

Model checking procedures

Choices of $g(u, x, X_i, \widehat{\beta})$

- functional form: $I(X_{iq} \leqslant x)$
- proportional means assumption: X_{iq}
- exponential link function: $I(\widehat{\beta}^{\mathrm{T}} X_i \leqslant x)$
- overall fit: $I(X_i \leqslant x)$

Graphical inspection

- plot $W(\infty, x)$ and a few realizations from $\widetilde{W}(\infty, x)$ against x
- plot $W(t, \cdot)$ and a few realizations from $\widetilde{W}(t, \cdot)$ against t

Model checking procedures

Choices of $g(u, x, X_i, \widehat{\beta})$

- functional form: $I(X_{iq} \leqslant x)$
- proportional means assumption: X_{iq}
- exponential link function: $I(\widehat{\beta}^{\mathrm{T}} X_i \leqslant x)$
- overall fit: $I(X_i \leqslant x)$

Graphical inspection

- plot $W(\infty, x)$ and a few realizations from $\widetilde{W}(\infty, x)$ against x
- plot $W(t, \cdot)$ and a few realizations from $\widetilde{W}(t, \cdot)$ against t

Supremum test

- generate a large number of, say 10000, realizations from $\sup_x |\widetilde{W}(\infty, x)|$
- compare them to the observed value of $\sup_x |W(\infty, x)|$

Outline

Simulation studies

- Intensity functions: $0.7(1+0.7t)^{-1}\eta\exp(\beta_{11}X_1+\beta_{12}X_2)$ and $0.4\eta\exp(\beta_{21}X_1+\beta_{22}X_2)$
- $(\beta_{11},\beta_{12})=(0.5,-0.5),(\beta_{21},\beta_{22})=(0,0.6)$
- $X_1\sim\mathrm{Ber}(0.5)$ and $X_2\sim\mathrm{Un}(0,1)$
- $\eta\mid X_1,X_2\sim\text{Gamma}(\text{mean}=1,\text{variance}=X_1+X_2)$
- Each subject has up to 3 examination times, uniformly distributed on $[0,3]$

Simulation studies

Table 1

Summary statistics for the simulation studies on bivariate panel count data

		Marginal model						Random-effects model			
n	Parameter	Bias	SE	SEE	CP	SEEn	CPn	Bias	SE	SEE	CP
200	$\beta_{11} = 0.5$	−0.003	0.218	0.232	95.9	0.129	75.4	−0.174	0.201	0.214	89.0
	$\beta_{12} = -0.5$	−0.011	0.398	0.398	94.6	0.199	66.8	−0.179	0.350	0.374	94.2
	$\beta_{21} = 0$	−0.007	0.225	0.212	94.5	0.100	63.7	−0.176	0.187	0.197	87.5
	$\beta_{22} = 0.6$	−0.006	0.374	0.408	96.4	0.179	65.1	−0.173	0.329	0.342	93.0
400	$\beta_{11} = 0.5$	−0.001	0.151	0.158	95.9	0.087	74.5	−0.175	0.141	0.149	80.1
	$\beta_{12} = -0.5$	−0.006	0.279	0.274	94.3	0.135	65.2	−0.173	0.247	0.258	91.2
	$\beta_{21} = 0$	−0.003	0.151	0.149	94.5	0.068	62.3	−0.178	0.130	0.138	76.6
	$\beta_{22} = 0.6$	−0.005	0.263	0.275	95.6	0.121	63.3	−0.171	0.227	0.237	89.9
800	$\beta_{11} = 0.5$	0.002	0.107	0.110	95.4	0.060	73.1	−0.175	0.100	0.105	62.0
	$\beta_{12} = -0.5$	−0.003	0.196	0.193	94.7	0.093	64.7	−0.175	0.172	0.179	85.1
	$\beta_{21} = 0$	−0.001	0.106	0.105	94.7	0.046	61.2	−0.177	0.093	0.097	55.9
	$\beta_{22} = 0.6$	−0.001	0.184	0.189	95.4	0.083	61.8	−0.174	0.160	0.165	82.7

Note: SE, SEE, and CP denote standard error, mean standard error estimator, and coverage probability of the 95% confidence interval. SEEn and CPn denote the mean standard error estimator and coverage of the 95% confidence interval by the naive method ignoring dependence of related event times.

Simulation studies

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

Table 1

Summary statistics for the simulation studies on bivariate panel count data

		Marginal model						Random-effects model			
n	Parameter	Bias	SE	SEE	CP	SEEn	CPn	Bias	SE	SEE	CP
200	$\beta_{11} = 0.5$	−0.003	0.218	0.232	95.9	0.129	75.4	−0.174	0.201	0.214	89.0
	$\beta_{12} = -0.5$	−0.011	0.398	0.398	94.6	0.199	66.8	−0.179	0.350	0.374	94.2
	$\beta_{21} = 0$	−0.007	0.225	0.212	94.5	0.100	63.7	−0.176	0.187	0.197	87.5
	$\beta_{22} = 0.6$	−0.006	0.374	0.408	96.4	0.179	65.1	−0.173	0.329	0.342	93.0
400	$\beta_{11} = 0.5$	−0.001	0.151	0.158	95.9	0.087	74.5	−0.175	0.141	0.149	80.1
	$\beta_{12} = -0.5$	−0.006	0.279	0.274	94.3	0.135	65.2	−0.173	0.247	0.258	91.2
	$\beta_{21} = 0$	−0.003	0.151	0.149	94.5	0.068	62.3	−0.178	0.130	0.138	76.6
	$\beta_{22} = 0.6$	−0.005	0.263	0.275	95.6	0.121	63.3	−0.171	0.227	0.237	89.9
800	$\beta_{11} = 0.5$	0.002	0.107	0.110	95.4	0.060	73.1	−0.175	0.100	0.105	62.0
	$\beta_{12} = -0.5$	−0.003	0.196	0.193	94.7	0.093	64.7	−0.175	0.172	0.179	85.1
	$\beta_{21} = 0$	−0.001	0.106	0.105	94.7	0.046	61.2	−0.177	0.093	0.097	55.9
	$\beta_{22} = 0.6$	−0.001	0.184	0.189	95.4	0.083	61.8	−0.174	0.160	0.165	82.7

Note: SE, SEE, and CP denote standard error, mean standard error estimator, and coverage probability of the 95% confidence interval. SEEn and CPn denote the mean standard error estimator and coverage of the 95% confidence interval by the naive method ignoring dependence of related event times.

Simulation studies

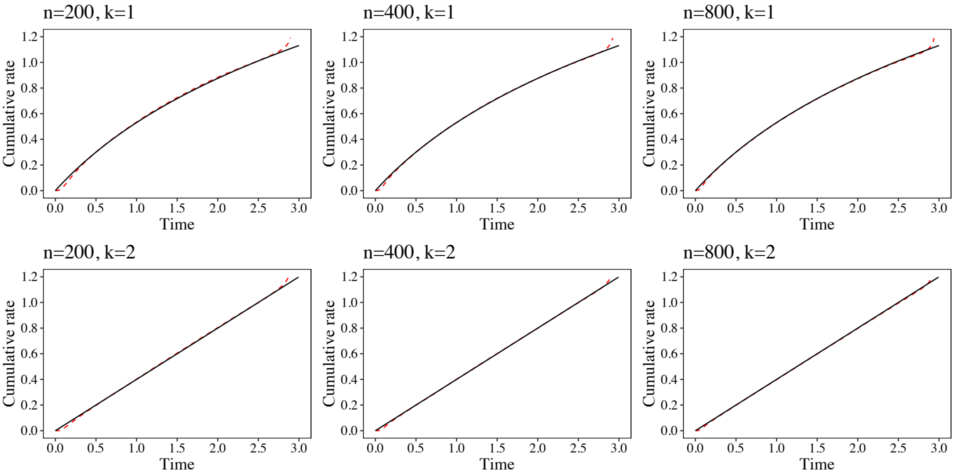
Table 1

Summary statistics for the simulation studies on bivariate panel count data

		Marginal model						Random-effects model			
n	Parameter	Bias	SE	SEE	CP	SEEn	CPn	Bias	SE	SEE	CP
200	$\beta_{11} = 0.5$	−0.003	0.218	0.232	95.9	0.129	75.4	−0.174	0.201	0.214	89.0
	$\beta_{12} = -0.5$	−0.011	0.398	0.398	94.6	0.199	66.8	−0.179	0.350	0.374	94.2
	$\beta_{21} = 0$	−0.007	0.225	0.212	94.5	0.100	63.7	−0.176	0.187	0.197	87.5
	$\beta_{22} = 0.6$	−0.006	0.374	0.408	96.4	0.179	65.1	−0.173	0.329	0.342	93.0
400	$\beta_{11} = 0.5$	−0.001	0.151	0.158	95.9	0.087	74.5	−0.175	0.141	0.149	80.1
	$\beta_{12} = -0.5$	−0.006	0.279	0.274	94.3	0.135	65.2	−0.173	0.247	0.258	91.2
	$\beta_{21} = 0$	−0.003	0.151	0.149	94.5	0.068	62.3	−0.178	0.130	0.138	76.6
	$\beta_{22} = 0.6$	−0.005	0.263	0.275	95.6	0.121	63.3	−0.171	0.227	0.237	89.9
800	$\beta_{11} = 0.5$	0.002	0.107	0.110	95.4	0.060	73.1	−0.175	0.100	0.105	62.0
	$\beta_{12} = -0.5$	−0.003	0.196	0.193	94.7	0.093	64.7	−0.175	0.172	0.179	85.1
	$\beta_{21} = 0$	−0.001	0.106	0.105	94.7	0.046	61.2	−0.177	0.093	0.097	55.9
	$\beta_{22} = 0.6$	−0.001	0.184	0.189	95.4	0.083	61.8	−0.174	0.160	0.165	82.7

Note: SE, SEE, and CP denote standard error, mean standard error estimator, and coverage probability of the 95% confidence interval. SEEn and CPn denote the mean standard error estimator and coverage of the 95% confidence interval by the naive method ignoring dependence of related event times.

Simulation studies



Simulation studies

Table 2
Type I error rates for different types of supremum tests

	First event			Second event		
Test	$n = 200$	$n = 400$	$n = 800$	$n = 200$	$n = 400$	$n = 800$
Proportionality	0.037	0.044	0.046	0.038	0.042	0.047
Functional form	0.027	0.037	0.043	0.029	0.038	0.042
Link function	0.029	0.039	0.042	0.032	0.039	0.042
Omnibus	0.021	0.030	0.037	0.020	0.031	0.042

Outline

A skin cancer trial

Basal cell carcinoma and squamous cell carcinoma

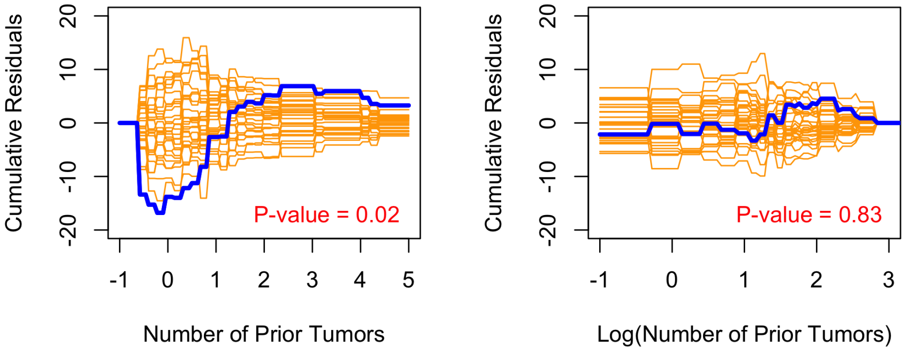
- 143 patients were randomized to receive treatment
- 147 were assigned to placebo

Examinations: every 6 months

Covariates:

- treatment indicator
- gender
- age at diagnosis dichotomized as $\geqslant 65$ versus < 65 years
- **number of prior skin tumors at baseline**

A skin cancer trial



A skin cancer trial

Table 4

Regression analysis of panel count data in a skin cancer chemoprevention trial

	Basal cell carcinoma			Squamous cell carcinoma			Any cancer		
	Estimate	Std error	p-value	Estimate	Std error	p-value	Estimate	Std error	p-value
Marginal model									
Treatment	−0.167	0.152	0.274	−0.008	0.273	0.976	−0.108	0.138	0.436
Log(prior tumors)	0.730	0.083	$< 10^{-3}$	0.927	0.159	$< 10^{-3}$	0.791	0.083	$< 10^{-3}$
Male	0.045	0.184	0.806	0.560	0.380	0.141	0.209	0.163	0.200
Age$\geqslant$65	−0.210	0.154	0.172	0.741	0.262	0.005	0.111	0.132	0.398
Random-effects model									
Treatment	−0.095	0.173	0.582	−0.137	0.188	0.465	−0.056	0.141	0.692
Log(prior tumors)	0.861	0.105	$< 10^{-3}$	0.903	0.120	$< 10^{-3}$	0.874	0.084	$< 10^{-3}$
Male	0.137	0.170	0.419	0.669	0.218	0.002	0.272	0.147	0.064
Age$\geqslant$65	−0.236	0.178	0.183	0.808	0.251	0.001	0.097	0.154	0.530

Thank you!