Transitional Analysis of Longitudinal Data

Leilei Zeng

lzeng@sfu.ca

Department of Statistics & Actuarial Science

Simon Fraser University

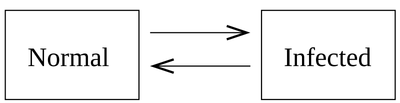
Oct. 14, 2005

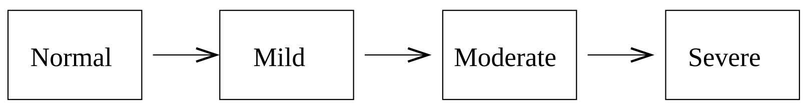
OUTLINE

- Motivation
- Multivariate/Clustered Multi-State Processes
  - Marginal Methods Based Estimating Functions
  - Simulation, Application and Extensions
  - Design Criteria
- Open Problems in Transitional Analysis
  - Incomplete Multi-State Data
  - Mixed Transitional Models

MOTIVATION

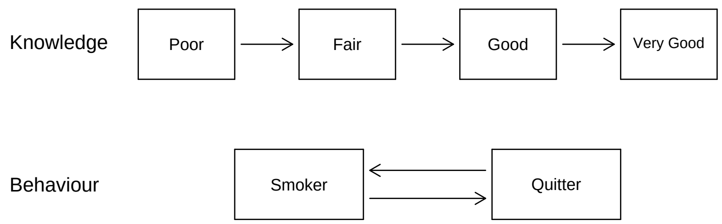
- Longitudinal studies :
  $\Longrightarrow$ measurements are taken at a set of planned visits
- Frameworks for Longitudinal Data Analysis (Neuhaus, 1992)
  - Marginal Methods
  - Hierarchical Random-effects Models
  - Transition Models
- Transitional models are appropriate when interest lies in characterizing rates of change in dynamic processes



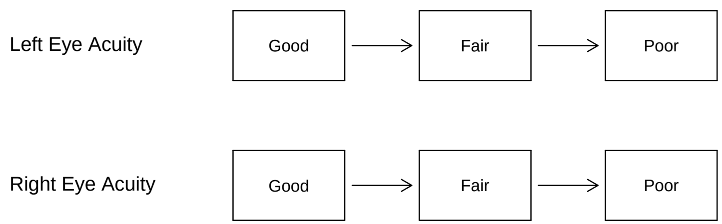


MOTIVATION

- Interesting problems exist involving two or more correlated processes
  - studying effects of health promotion interventions on knowledge, attitudes, and behavior



  - examining deterioration of paired organ systems (opthalmology, nephrology, etc.)



  - when assessing observer agreement on dynamic disease processes

EXAMPLE I

Indonesia Children’s Study (Zeger and Liang, 1991)

- 3000 pre-school children were medically examined for up to 7 visits
- presence/absence of xerophthalmia and respiratory infection, and other subject characteristics are determined

Features:

- Children with xerophthalmia are more likely to suffer respiratory infections which in turn deplete stores of Vitamin A and increase the risk of developing xerophthalmia
- Correlated Multivariate Processes

EXAMPLE II

Behavioral Studies of Childhood Smoking

- 100 elementary schools were randomized to an intensive or standard anti-smoking program
- smoking behavior, reasons for smoking, attitudes towards smoking are assessed annually from grade 6 to grade 12

Objective:

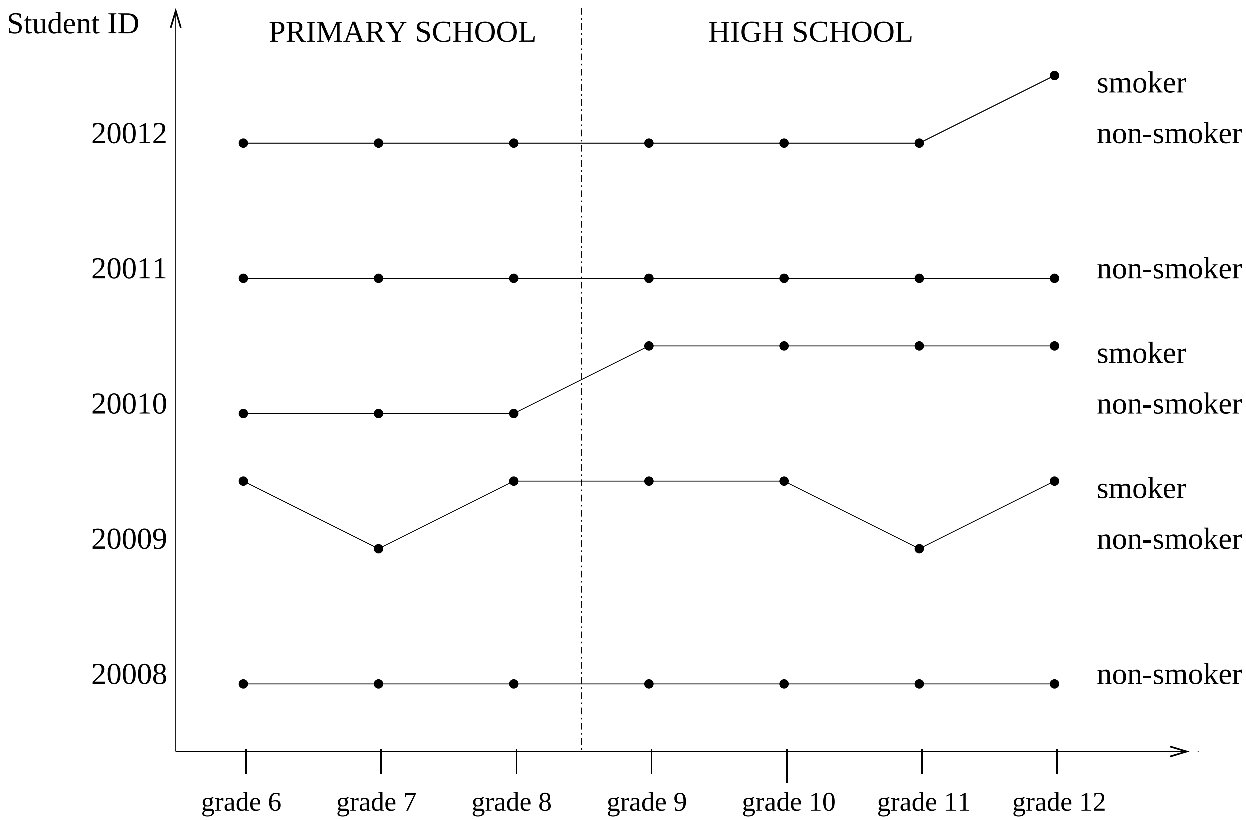
- smoking behavior change (e.g. non-smoking to smoking)
- effectiveness of intensive program

Features:

- Cluster-randomized Trials

EXAMPLE II

Behavioral Studies of Smoking (some students from school 1)



EXAMPLE III

A Study in Intensive Care Unit

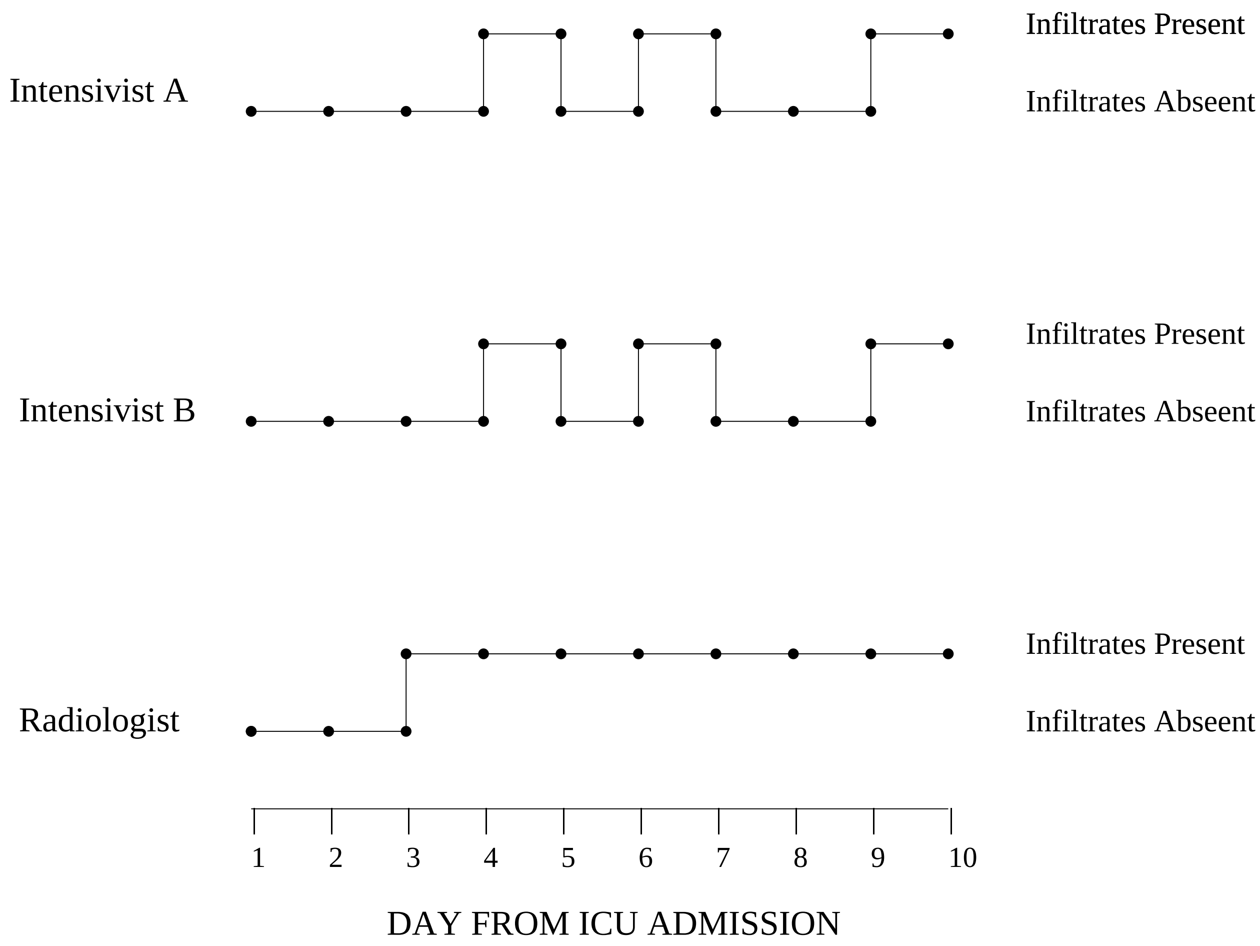
- 120 adults with acute lung injury from ICU were scheduled to have daily chest radiographs
- 3 raters interpreted each radiograph independently to assess the presence of diffuse bilateral infiltrates

Objective:

- Accuracy and reliability of the classification of diffuse bilateral infiltrates
- Assess changes of a rater's interpretations over time, and how these changes are associated with other rater's changes in interpretation

EXAMPLE III

ICU study (10 films from a particular patient)



MOTIVATION

Correlated Multi-state Processes, interest may lie in

- characterizing how these processes change together
- estimating or testing treatment effects on rates of change
- improving efficiency through joint modeling

Options for joint analysis

- random effect models
- predicting one process from another

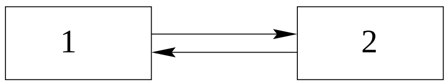
Focus: Marginal methods based on estimating functions

- retains marginal interpretation of covariate effects
- associations modeled via second order estimating functions
- robust “sandwich” variance estimates

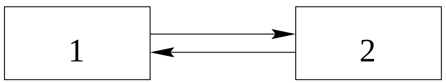
MULTIPLE MULTI-STATE PROCESSES

For simplicity, we focus on J two-state processes

multivariate longitudinal binary data



Process 1



Process 2

Notation

- N individuals
- $t_1, \ldots, t_K$ scheduled assessment times
- J binary responses measured longitudinally
- $Y_i^{(j)}(t_k) = 1, 2$, state occupied by subject i at t_k for process j
- $\boldsymbol{x}_i^{(j)}(t_k) = (1, x_{i1}^{(j)}(t_k), \ldots, x_{i,p_j-1}^{(j)}(t_k))'$ and $\boldsymbol{x}_i(t_k) = (\boldsymbol{x}_i^{(j)}(\boldsymbol{t_k})', j = 1, \cdots, J)'$

MARGINAL TRANSITION MODELS

Illustrative Assumptions

- Assume a first order Markov chain for each response process

$$\text{Pr}(Y^{(j)}(t_{k+1}) \mid Y^{(j)}(t_k), \ldots, Y^{(j)}(t_1), \boldsymbol{x}(t_k)) = \text{Pr}(Y^{(j)}(t_{k+1}) \mid Y^{(j)}(t_k), \boldsymbol{x}(t_k))$$

- Transition probabilities

$$\pi_\ell^{(j)}(t_k) = P(Y^{(j)}(t_{k+1}) = 3 - \ell \mid Y^{(j)}(t_k) = \ell, \boldsymbol{x}(t_k))$$

- Models for the transition probability

$$\text{logit}(\pi_\ell^{(j)}(t_k)) = \boldsymbol{x}^{(j)}(t_k)' \boldsymbol{\beta}_\ell^{(j)}$$

  - $\boldsymbol{\beta}^{(j)} = (\boldsymbol{\beta}_1^{(j)'}, \boldsymbol{\beta}_2^{(j)'})'$
  - $\boldsymbol{\beta} = (\boldsymbol{\beta}^{(1)'}, \ldots, \boldsymbol{\beta}^{(J)'})'$

TRANSITIONAL DATA

Data For First Moments

- Define event of occurrence of $\ell \to 3 - \ell$ transition
  - “At risk" indicator: $\delta_\ell^{(j)}(t_k) = I(Y^{(j)}(t_k) = \ell)$
  - $\boldsymbol{\delta}(t_k) = (\delta_1^{(1)}(t_k), \delta_2^{(1)}(t_k), \ldots, \delta_1^{(J)}(t_k), \delta_2^{(J)}(t_k))$
  - Transition indicator: $N_\ell^{(j)}(t_k) = I(Y^{(j)}(t_{k+1}) = 3 - \ell,\ \ Y^{(j)}(t_k) = \ell)$
- Then under a first-order Markov chain

$$\mathsf{E}(N_\ell^{(j)}(t_k) \mid \delta_\ell^{(j)}(t_k) = 1, \boldsymbol{x}(t_k)) = \pi_\ell^{(j)}(t_k),$$

- Finally

$$N^{(j)}(t_k) = \sum_{\ell=1,2} \delta_\ell^{(j)}(t_k) N_\ell^{(j)}(t_k) \qquad \pi^{(j)}(t_k) = \sum_{\ell=1,2} \delta_\ell^{(j)}(t_k) \pi_\ell^{(j)}(t_k)$$

ASSOCIATION STRUCTURE

Data For Second Moments

		$N_\ell^{(1)}(t_k)$ (Process 1)	
		1	0
$N_{\ell'}^{(2)}(t_k)$ (Process 2)	1	p_{11}	p_{01}
	0	p_{10}	p_{00}

- Define conditional odds

$$ODDS(N_\ell^{(j)}(t_k) \mid N_{\ell'}^{(j')}(t_k)) = \frac{P(N_\ell^{(j)}(t_k)=1 \mid N_{\ell'}^{(j')}(t_k), \delta_\ell^{(j)}(t_k)=\delta_{\ell'}^{(j')}(t_k)=1, \boldsymbol{x}(t_k))}{1-P(N_\ell^{(j)}(t_k)=1 \mid N_{\ell'}^{(j')}(t_k), \delta_\ell^{(j)}(t_k)=\delta_{\ell'}^{(j')}(t_k)=1, \boldsymbol{x}(t_k))}$$

- The associations are measured by pairwise odds ratio:

$$\gamma_{\ell\ell'}^{(jj')}(t_k) = \frac{ODDS(N_\ell^{(j)}(t_k) \mid N_{\ell'}^{(j')}(t_k) = 1)}{ODDS(N_\ell^{(j)}(t_k) \mid N_{\ell'}^{(j')}(t_k) = 0)}$$

MODELS FOR ASSOCIATION

- Regression models:

$$\log(\gamma_{\ell\ell'}^{(jj')}(t_k)) = \boldsymbol{w}_{\ell\ell'}^{(jj')}(t_k)'\boldsymbol{\psi}$$

  - $\boldsymbol{w}_{\ell\ell'}^{(jj')}(t_k)$ is a vector of explanatory variables
  - $\boldsymbol{\psi}$ is a vector of association parameters

- A particular regression model can be

$$\log(\gamma_{\ell\ell'}^{(jj')}(t_k)) = \psi_0 + \psi_1 I(\ell = 1, \ell' = 2) + \psi_2 I(\ell = 2, \ell' = 1) + \psi_3 I(\ell = 2, \ell' = 2)$$

  - association among transitions in the "same" direction:
    $\gamma_{11}^{(jj')}(t_k) = \exp(\psi_0), \quad \gamma_{22}^{(jj')}(t_k) = \exp(\psi_0 + \psi_3)$
  - association among transitions in the "opposite" direction:
    $\gamma_{12}^{(jj')}(t_k) = \exp(\psi_0 + \psi_1), \quad \gamma_{21}^{(jj')}(t_k) = \exp(\psi_0 + \psi_2)$

MODELS FOR ASSOCIATION

- Expectations for pairwise products of transition indicators

$$\mathsf{E}\left(N^{(j)}(t_k)N^{(j')}(t_k) \mid \boldsymbol{\delta}(t_k), \boldsymbol{x}(t_k)\right) = \sum_{\ell=1,2} \delta_\ell^{(j)}(t_k)\delta_{\ell'}^{(j')}(t_k)\pi_{\ell\ell'}^{(jj')}(t_k)$$

  - $\pi_{\ell\ell'}^{(jj')}(t_k) = \mathsf{Pr}(N_\ell^{(j)}(t_k) = 1, N_{\ell'}^{(j')}(t_k) = 1 \mid \delta_\ell^{(j)}(t_k) = 1, \delta_{\ell'}^{(j')}(t_k) = 1)$
  - $\pi_{\ell\ell'}^{(jj')}(t_k)$ is a function of $\pi_\ell^{(j)}(t_k)$, $\pi_{\ell'}^{(j')}(t_k)$, and $\gamma_{\ell\ell'}^{(jj')}(t_k)$

- Option for Joint estimation
  - GEE2 type of methods (Prentice, 1988)
  - Alternating Logistic Regression (ALR) methods (Carey et al., 1993)

ESTIMATION OF β

- Estimating equations for regression parameters $\boldsymbol{\beta}$

$$U_1(\boldsymbol{\beta}, \boldsymbol{\psi}) = \sum_i \sum_{k=1}^{K-1} D(t_k; \boldsymbol{\beta})' V(t_k; \boldsymbol{\beta}, \boldsymbol{\psi})^{-1} (\boldsymbol{N}(t_k) - \boldsymbol{\pi}(t_k)) = 0 \qquad (1)$$

  - $\boldsymbol{N}(t_k) = (N^{(1)}(t_k), \ldots, N^{(J)}(t_k))'$
  - $\boldsymbol{\pi}(t_k) = (\pi^{(1)}(t_k), \ldots, \pi^{(J)}(t_k))'$
  - $V(t_k; \boldsymbol{\beta}, \boldsymbol{\phi})$ is a covariance matrix of $\boldsymbol{N}(t_k)$
  - $D(t_k; \boldsymbol{\beta}) = \partial \boldsymbol{\pi}(t_k) / \partial \boldsymbol{\beta}'$
  - Consistent estimate of $\boldsymbol{\beta}$ can be obtained

$$\mathsf{E}(N^{(j)}(t_k) - \pi^{(j)}(t_k)) = \mathsf{E}_{\boldsymbol{\delta}(t_k)} \left(\mathsf{E}_{N^{(j)}(t_k)|\boldsymbol{\delta}(t_k)} (N^{(j)}(t_k) - \pi^{(j)}(t_k)) \right)$$

JOINT ESTIMATION

- ALR estimating equations for association parameter ψ

$$U_2(\boldsymbol{\beta}, \boldsymbol{\psi}) = \sum_i \sum_{k=1}^{K-1} C(t_k \mid \boldsymbol{\beta}, \boldsymbol{\psi})' S(t_k \mid \boldsymbol{\beta}, \boldsymbol{\psi})^{-1} \boldsymbol{\epsilon}(t_k) = 0. \quad (2)$$

  - $\xi^{(jj')}(t_k) = \mathsf{E}(N^{(j)}(t_k) \mid N^{(j')}(t_k), \boldsymbol{\delta}(t_k), \boldsymbol{x}(t_k))$
  - $\boldsymbol{\epsilon}(t_k) = (\epsilon^{(jj')}(t_k), j < j')'$ with $\epsilon^{(jj')}(t_k) = N^{(j)}(t_k) - \xi^{(jj')}(t_k)$
  - $C(t_k \mid \boldsymbol{\beta}, \boldsymbol{\psi}) = \partial \boldsymbol{\xi}(t_k) / \partial \boldsymbol{\psi}'$
  - $S(t_k \mid \boldsymbol{\beta}, \boldsymbol{\psi}) = \mathsf{diag}\{\xi^{(jj')}(t_k)(1 - \xi^{(jj')}(t_k)), j < j'\}$
- Solve $U_1(\boldsymbol{\beta}, \boldsymbol{\psi})$ and $U_2(\boldsymbol{\beta}, \boldsymbol{\psi})$ in an alternating fashion
- Model formulation can be generalized to deal with
  - Higher-Order Markov Processes
  - Multi-State Markov Processes

JOINT ESTIMATION

- GEE2 type of estimating functions for $\boldsymbol{\zeta} = (\boldsymbol{\beta}', \boldsymbol{\psi}')'$

$$\sum_i \sum_{k=1}^{K-1} \left(\frac{\partial(\boldsymbol{\pi}(t_k)', \boldsymbol{\eta}(t_k)')}{\partial \boldsymbol{\zeta}} \right) B^{-1}(t_k) \left(\begin{array}{c} \boldsymbol{N}(t_k) - \boldsymbol{\pi}(t_k) \\ \boldsymbol{Z}(t_k) - \boldsymbol{\eta}(t_k) \end{array} \right) = 0,$$

  - pairwise products $Z^{(jj')}(t_k) = N^{(j)}(t_k) N^{(j')}(t_k)$, $j < j'$
  - $\eta^{(jj')} = \mathsf{E}(Z^{(jj')}(t_k) \mid \boldsymbol{\delta}(t_k), \boldsymbol{x}(t_k))$
  - $\boldsymbol{Z}(t_k) = (Z^{(jj')}(t_k)', j < j')'$ and $\boldsymbol{\eta}(t_k) = (\eta^{(jj')}(t_k)', j < j')'$
  - $B(t_k) = COV(\boldsymbol{N}(t_k), \boldsymbol{Z}(t_k) \mid \boldsymbol{\delta}(t_k), \boldsymbol{x}(t_k))$
  - Involves higher-order moments of $\boldsymbol{N}(t_k)$
  - Replace with “working” covariance matrix

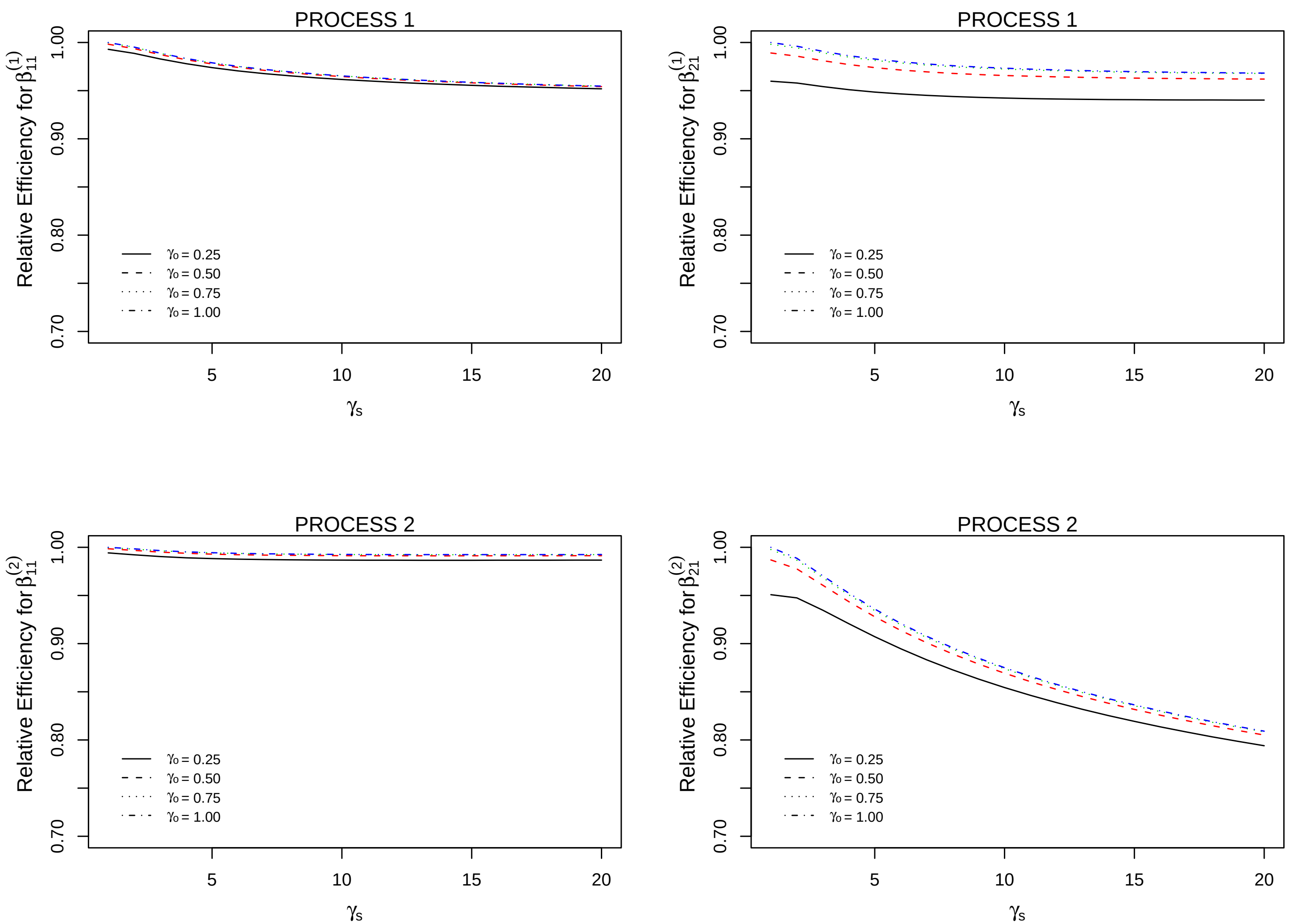
SIMULATION STUDIES

- Design (Indonesian Childrens Study)
  - Two processes: absence/presence of xerophthalmia, respiratory infection
  - $K = 7$ assessments
  - $N = 1000$ subjects
- Model configuration

$$\text{logit}(\pi_\ell^{(j)}(t_k)) = \beta_{\ell 0}^{(j)} + \beta_{\ell 1}^{(j)} x_{i1} \ , \qquad j = 1, 2, \ \ \ell = 1, 2,$$

  - $x_{i1} = 1$ if treated, $x_{i1} = 0$ otherwise
  - $\log(\gamma_{\ell\ell'}^{(jj')}(t_k)) = \psi_0 + \psi_1 I(\ell = 1, \ell' = 2) + \psi_2 I(\ell = 2, \ell' = 1) + \psi_3 I(\ell = 2, \ell' = 2)$
- Results
  - empirical biases are negligible
  - empirical coverage probabilities $\approx 95\%$

Asymptotic Relative Efficiency: Joint V.S. Separate



Process	baseline rates		treatment	
	$1 \to 2$	$2 \to 1$	$1 \to 2$	$2 \to 1$
1	50%	50%	logit(0.75)	logit(1.33)
2	25%	50%	logit(0.75)	logit(1.33)

APPLICATION

- Analysis
  - Assume two-state Markov process for each rater's response (first-order reasonable)

$$Y_i^{(j)}(t_k) = \begin{cases} 1, & \text{infiltrates present according to rater } j \\ 2, & \text{infiltrates not present according to rater } j \end{cases}$$

  - PPO/FCIO (lowest ratio of Partial Pressure of Oxygen to Fractional Concentration of Inspired Oxygen)
  - VENT (whether they had previous ventilation or not)
- Measure of Agreement in Change:
  - $\tau_{\ell m} = (\sqrt{\gamma_{\ell m}} - 1)/(\sqrt{\gamma_{\ell m}} + 1)$
  - $\widehat{\tau}_{11} = 0.66, 95\%$ CI$(0.57, 0.73)$; $\widehat{\tau}_{22} = 0.74, 95\%$ CI$(0.59, 0.83)$
  - $\widehat{\tau}_{12} = -0.60, 95\%$ CI$(-0.86, -0.11)$; $\widehat{\tau}_{21} = -0.56, 95\%$ CI$(-0.82, -0.10)$

APPLICATION

Table 1: Results from analyses of the onset and resolution of diffuse bilateral infiltrates in intensive care (Meade et al., 2000)

			Separate Analysis			Joint Analysis		
Raters	Transitions	Covariates	est.	s.e.	p-value	est.	s.e.	p-value
1	$1 \to 2$	Intercept	-1.153	0.294	< 0.001	-0.886	0.250	< 0.001
		VENT	-1.426	0.538	0.008	-1.545	0.485	0.002
		PPO/FCIO	-0.003	0.002	0.116	-0.004	0.002	0.024
	$2 \to 1$	Intercept	-1.357	0.409	0.001	-0.916	0.358	0.011
		VENT	-1.064	0.472	0.024	-0.871	0.360	0.016
		PPO/FCIO	0.005	0.003	0.077	0.004	0.002	0.136
2	$1 \to 2$	Intercept	-0.881	0.292	0.003	-0.616	0.263	0.019
		VENT	-1.302	0.490	0.008	-1.407	0.447	0.002
		PPO/FCIO	-0.005	0.002	0.015	-0.006	0.002	0.001
	$2 \to 1$	Intercept	-1.305	0.416	0.002	-0.836	0.371	0.024
		VENT	-1.012	0.444	0.023	-0.682	0.359	0.057
		PPO/FCIO	0.005	0.003	0.064	0.003	0.003	0.195
3	$1 \to 2$	Intercept	-1.235	0.460	0.007	-0.270	0.331	0.414
		VENT	-1.725	0.766	0.024	-2.267	0.674	0.001
		PPO/FCIO	-0.005	0.003	0.052	-0.009	0.002	< 0.001
	$2 \to 1$	Intercept	-3.118	0.463	< 0.001	-2.872	0.435	< 0.001
		VENT	0.125	0.515	0.808	0.409	0.402	0.309
		PPO/FCIO	0.006	0.003	0.028	0.006	0.003	0.051
Associations	Transitions	Parameters	est.	s.e.	p-value	est.	s.e.	p-value
	$(1 \to 2, 1 \to 2)$	$\log(\gamma_{11})$	-	-	-	3.156	0.293	< 0.001
	$(1 \to 2, 2 \to 1)$	$\log(\gamma_{12})$	-	-	-	-2.780	1.191	0.020
	$(2 \to 1, 1 \to 2)$	$\log(\gamma_{21})$	-	-	-	-2.526	1.088	0.020
	$(2 \to 1, 2 \to 1)$	$\log(\gamma_{12})$	-	-	-	3.761	0.530	< 0.001

CLUSTER-RANDOMIZED TRIALS

- With clustered multi-state data
  - i indexes clusters, $i = 1, \ldots, I$
  - j indexes subjects
  - $\boldsymbol{x}_i^{(j)}(t_k) = \boldsymbol{x}_i(t_k)$, covariate vectors fixed over j
  - $\boldsymbol{\beta}_\ell^{(j)} = \boldsymbol{\beta}_\ell$, common covariate effects
  - Intra-cluster associations in terms of odds ratio $\gamma_{\ell\ell'}^{(jj')}(t_k)$
- Marginal methods for clustered multi-state longitudinal data
  - Estimating functions

$$U(\boldsymbol{\beta}, \boldsymbol{\psi}) = \sum_{i=1}^{I} U_i(\boldsymbol{\beta}, \boldsymbol{\psi} = \sum_{i=1}^{I} \begin{pmatrix} U_{1i}(\boldsymbol{\beta}, \boldsymbol{\psi}) \\ U_{2i}(\boldsymbol{\beta}, \boldsymbol{\psi}) \end{pmatrix}$$

  - Sample size criteria are based on model based variance estimates

DERIVATION OF SAMPLE SIZE CRITERIA

- Consider
  - Progressive Process ($1 \longrightarrow 2$ transition only)
  - Constant treatment effect (logit$(\pi_{i1}(t_k)) = \beta_{0k} + \beta_1 x_{i1}$)
  - Constant intra-cluster association ($\gamma_{\ell\ell'}^{(jj')}(t_k) = \gamma_{11}$)
- At the Design Stage
  - $H_0 : \beta_1 = \beta_0$ v.s $H_A : \beta_1 = \beta_A$ with level α_1 and power $1 - \alpha_2$
  - Formula for minimum required number of clusters

$$I \geq \left(\frac{Z_{\alpha_1/2}\sqrt{[\mathcal{I}(\boldsymbol{\zeta})^{-1}]_{K+1|\beta_1=\beta_0}} - Z_{1-\alpha_2}\sqrt{[\mathcal{I}(\boldsymbol{\zeta})^{-1}]_{K+1|\beta_1=\beta_A}}}{\beta_A - \beta_0} \right)^2$$

  - $\mathcal{I}(\boldsymbol{\zeta}) = \mathcal{I}(\boldsymbol{\beta}, \boldsymbol{\psi}) = \mathsf{E}(U_i(\boldsymbol{\beta}, \gamma)U_i(\boldsymbol{\beta}, \gamma)')$
  - $Z_\alpha = \Phi(1 - \alpha)^{-1}$ and $\Phi(\cdot)$ is the cumulative distribution function for a standard normal random variable

GENERAL REMARKS

- Multivariate Transition Models
  - Enable one to model associations between transitions from marginal models with marginal interpretations of covariate effects
  - Result in consistent estimates for both regression and association parameters
  - Increase efficiency compared to separate analyses
  - Can deal with categorical responses and higher order Markov processes
  - Flexible and can be used in a wide variety of settings

Open Problems in Transitional Analysis

Missing Data (monotone, and intermittent)

- Expanding the state space
  - joint models (Albert, 2000)
  - $Z_i^{(j)}(t_k) = Y_i^{(j)}(t_k)$ if observed, or $Z_i^{(j)}(t_k) = M + 1$ if missing
- Inverse Probability Weighted GEE (Robins et al., 1995)
- Missing (time-dependent) covariates

Variable assessment times

- continuous time processes
- intermittent assessments (panel data)

Mixed Transitional Models for Clustered Longitudinal Data