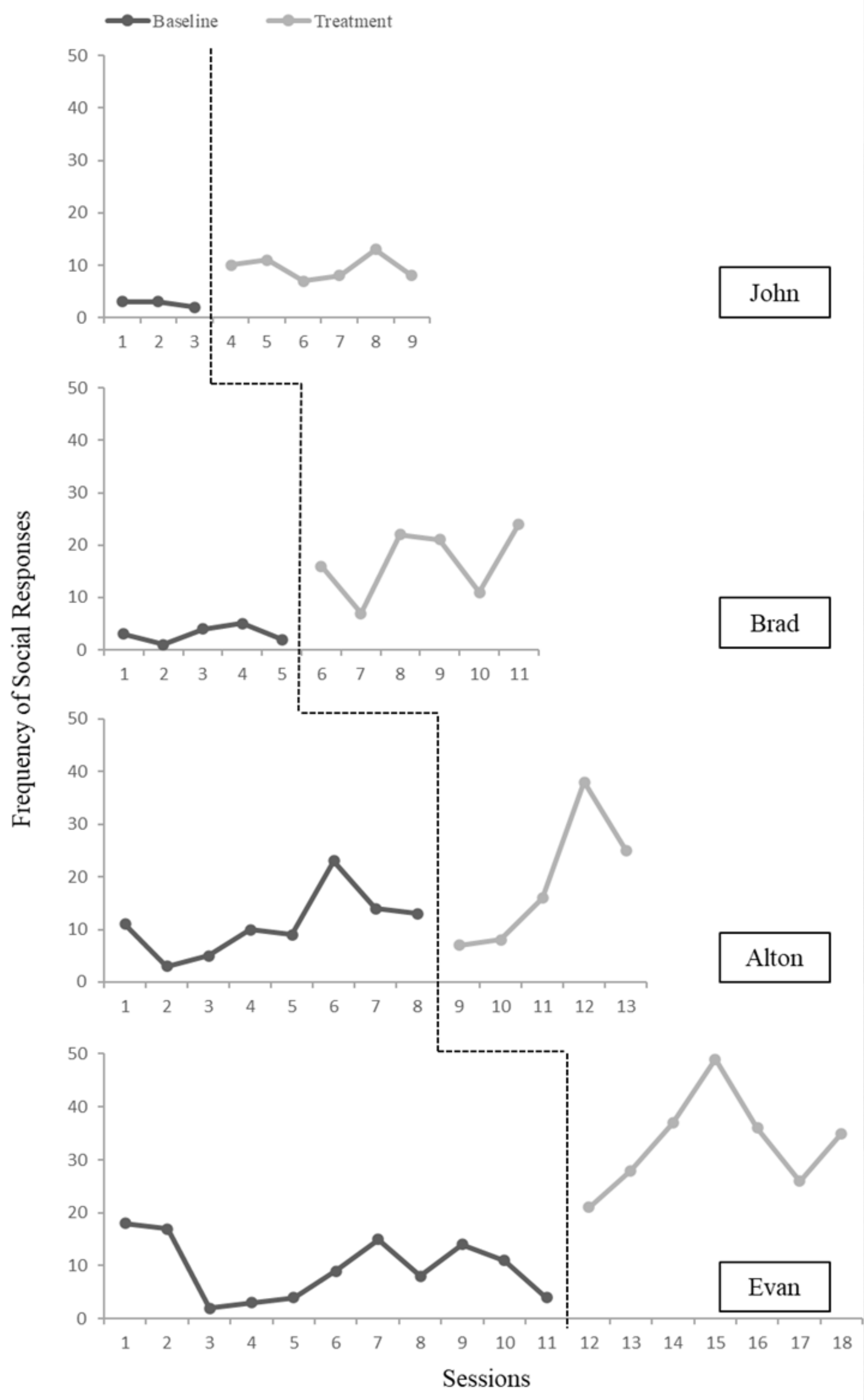
Model Selection of GLMMs in the Analysis of Count Data in SCEDs: A Monte Carlo Simulation

Haoran Li (haoranli@umn.edu)
University of Minnesota
June 26, 2024

Single-Case Research

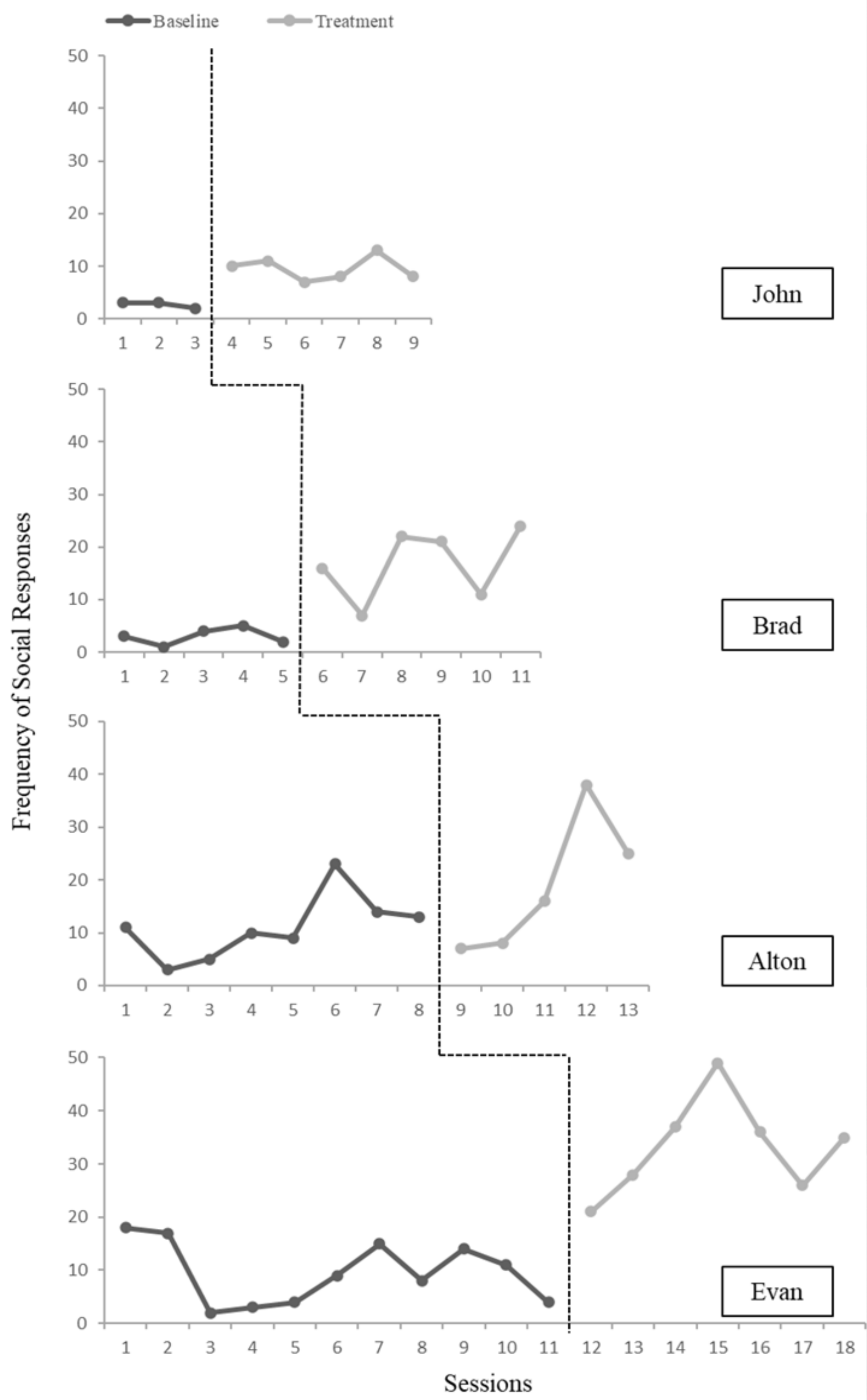
Single-case research is the intensive study of a **single case** or **small samples** by **repeatedly measuring** an outcome before and after an intervention to examine treatment effects.



Multiple Baseline Design

Multiple baseline design (MBD) is comprised of interrupted time series data from multiple cases, settings, or behaviors where an intervention is introduced **sequentially** within different time series (Baek & Ferron, 2013; Ferron et al., 2010).

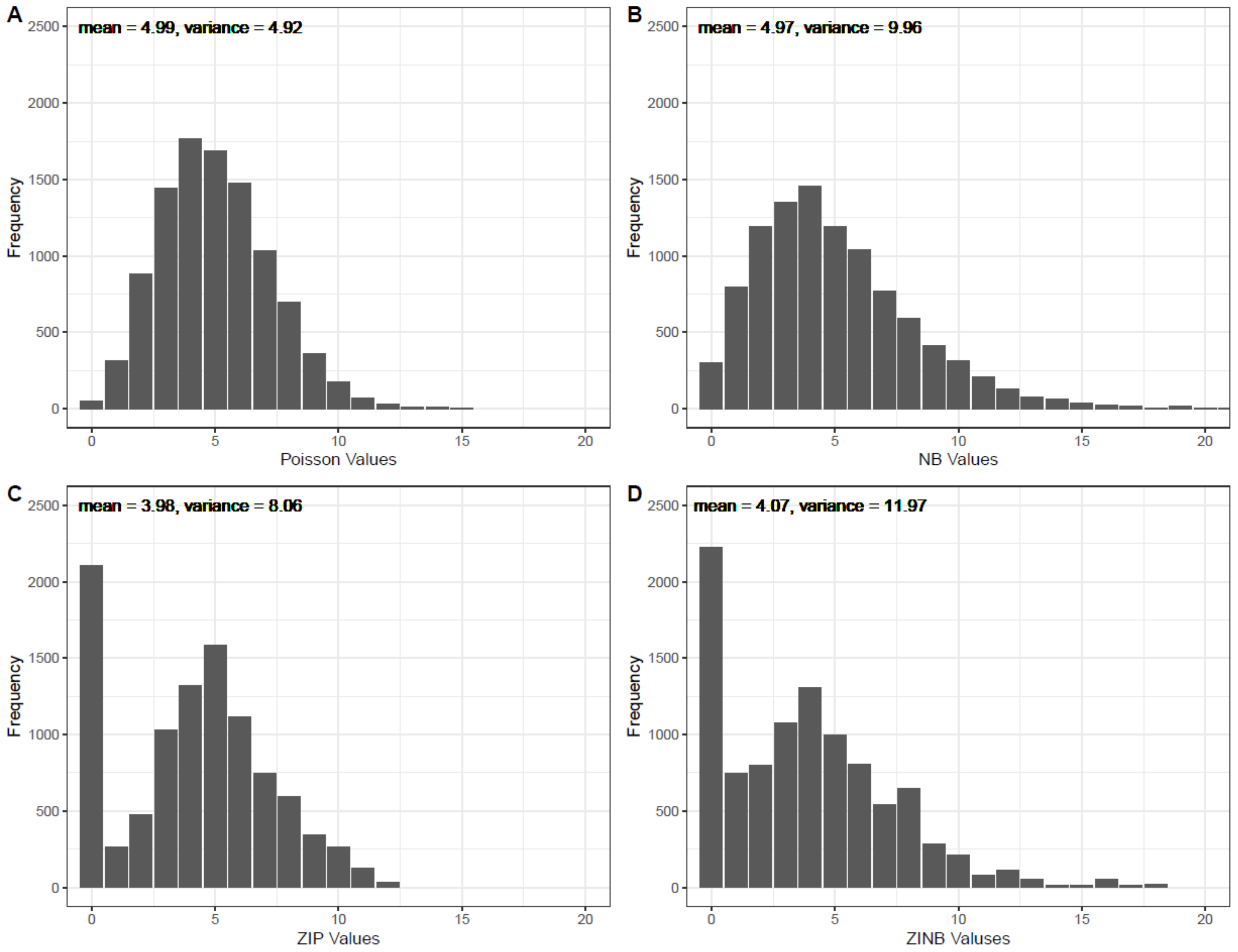
- The basic interrupted time series in MBD include two phases: baseline and treatment.
- Inferences about the intervention are usually made by comparing different conditions (baseline vs. treatment) presented to cases over time.



Challenges

- Nonnormal outcomes in SCEDs such as count and proportion data
- Autocorrelated count data with trend effects
- Overdispersion and Zero-inflation
- Model selection of optimal distributions

Count Distributions



GLMMs for SCEDs

$$Y_{ij} \sim Poisson\left(\lambda_{ij}\right)$$

Level 1: $\log(\lambda_{ij}) = \beta_{0j} + \beta_{1j} Phase_{ij}$

Level 2: $\begin{cases} \beta_{0j} = \gamma_{00} + u_{0j} \\ \beta_{1j} = \gamma_{10} + u_{1j} \end{cases}$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{u0}^2 & \sigma_{u0u1} \\ \sigma_{u1u0} & \sigma_{u1}^2 \end{bmatrix}\right)$$

$$Y_{ij} \sim Poisson\left(\lambda_{ij}\right)$$

Level 1: $\log(\lambda_{ij}) = \beta_{0j} + \beta_{1j} Time_{ij} + \beta_{2j} Phase_{ij} + \beta_{3j} Time'_{ij} Phase_{ij}$

Level 2: $\begin{cases} \beta_{0j} = \gamma_{00} + u_{0j} \\ \beta_{1j} = \gamma_{10} + u_{1j} \\ \beta_{2j} = \gamma_{20} + u_{2j} \\ \beta_{3j} = \gamma_{30} + u_{3j} \end{cases}$ $\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \\ u_{3j} \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{u0}^2 & \sigma_{u0u1} & \sigma_{u0u2} & \sigma_{u0u3} \\ \sigma_{u1u0} & \sigma_{u1}^2 & \sigma_{u1u2} & \sigma_{u1u3} \\ \sigma_{u2u0} & \sigma_{u2u1} & \sigma_{u2}^2 & \sigma_{u2u3} \\ \sigma_{u3u0} & \sigma_{u3u1} & \sigma_{u3u2} & \sigma_{u3}^2 \end{bmatrix}\right)$

The Poisson distribution assumes that the $E(Y) = Var(Y)$, which is often violated due to a data issue called overdispersion.

Overdispersion

Overdispersion in count data occurs when there is **excessive variance** than what a Poisson can deal with.

- Overdispersed count data: $Var(Y) > E(Y) = \lambda$

Overdispersion source: **correlated measurements**, **extra noise**, and **zero-inflation**.

Overdispersion is not uncommon for count data in SCEDs (Pustejovsky et al., 2019).

Ignoring overdispersion could lead to **biased** standard errors and **inflated** Type I error rates (Hilbe, 2011, 2014; Li et al., 2023).

Models to Handle Overdispersed Count Data

Negative binomial: $Y_{ij} \sim$ Negative binomial $\left(\lambda_{ij}, \theta\right)$

- $E(Y) = \lambda$ and $var(Y) = \lambda + \frac{\lambda^2}{\theta}, \theta > 0$

Observation-level random effects (OLRE) model: $Y_{ij} \sim$ Poisson (λ_{ij})

- $\log\left(\lambda_{ij}\right) = \beta_{0j} + \beta_{1j} Phase_{ij} + e_{ij}, e_{ij} \sim N(0, \sigma_e^2)$
- $E(Y) = \lambda$ and $var(Y) = \lambda + \lambda^2[\exp\left(\sigma_e^2\right) - 1]$

Li, H., Luo, W., Baek, E., Thompson, C. G., & Lam, K. (2023). Multilevel modeling in single-case studies with count and proportion data: A demonstration and evaluation. *Psychological Methods*. Advance online publication. https://doi.org/10.1037/met0000607

GLMMs with SCED count data

Performance

When count data are Overdispersed:

	Data generation	Fitted model	Estimation method	Estimates accurate?	Inferential results reliable?
Count data	Negative binomial	Poisson	Laplace (Wald test)	✓	X
		Negative binomial		✓	X
		OLRE		✓	X
	Negative binomial	Poisson	Pseudo likelihood (t test with Kenward-Roger)	✓	X
		Negative binomial		✓	✓
		OLRE		✓	✓
Count data	OLRE	Poisson	Laplace (Wald test)	✓	X
		Negative binomial		✓	X
		OLRE		✓	X
	OLRE	Poisson	Pseudo likelihood (t test with Kenward-Roger)	✓	X
		Negative binomial		✓	✓
		OLRE		✓	✓

Models to Handle Zero-Inflated Count Data

ZIP model:

Count component:

$$Y_{ij} \sim Poisson\ (\lambda_{ij})$$

$$\text{Level 1: } \log(\lambda_{ij}) = \beta_{0j} + \beta_{1j}Phase_{ij}$$

$$\text{Level 2: } \begin{cases} \beta_{0j} = \gamma_{00} + u_{0j} \\ \beta_{1j} = \gamma_{10} + u_{1j} \end{cases}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{u0} & \sigma_{u0u1} \\ \sigma_{u1u0} & \sigma^2_{u1} \end{bmatrix} \right)$$

Zero component:

$$logit(\pi_{ij}) = \log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_0 + \beta_1 * Phase_{ij}$$

ZINB model:

$$Y_{ij} \sim Negative\ Binomial\ (\lambda_{ij}, \theta)$$

- When data are zero-inflated, ZIP and ZINB models lead to more accurate estimates for the treatment effect ($\gamma 10$).
- When data are generated from Poisson and NB, ZIP and ZINB lead to more biased estimates for the treatment effect than Poisson and NB models.

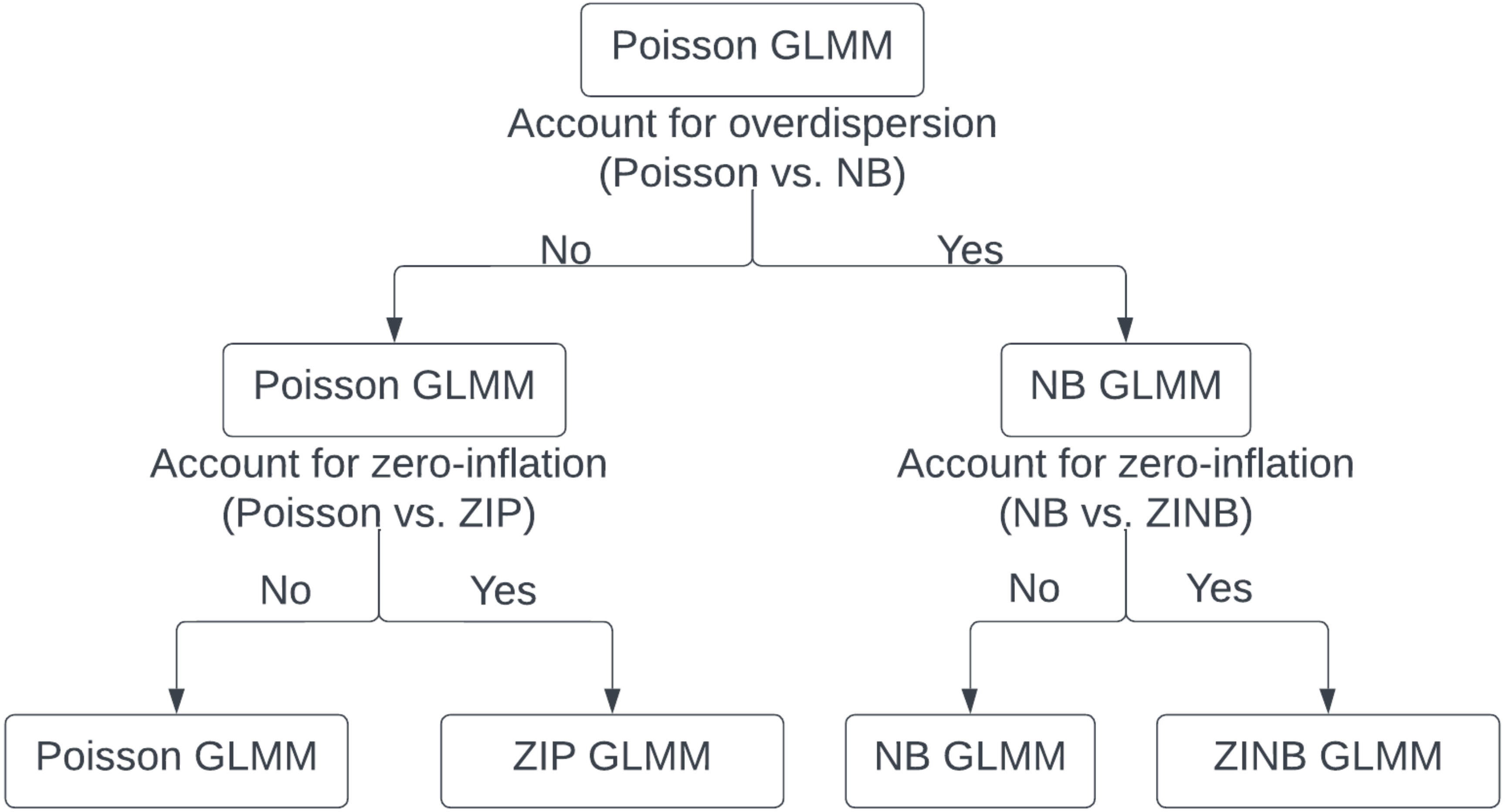
Li, H., Luo, W., & Baek, E. (2024). Multilevel modeling in single-case studies with zero-inflated and overdispersed count data. *Behavior Research Methods*. Advance online publication. https://doi-org/10.3758/s13428-024-02359-7

A Remaining Issue

- Most applied researchers are not aware of how to choose an appropriate distribution when dealing with SCED count data using GLMMs.

- Various model selection approaches

A Multi-stage Framework



Model Selection Strategies

Strategies	Overdispersion	Zero-inflation
Multi-stage		
Strategy 1	Pearson's chi-squared ($\alpha = .05$)	Parametric bootstrap ($\alpha = .05$)
Strategy 2	LRT ($\alpha = .05$)	Parametric bootstrap ($\alpha = .05$)
Strategy 3	Pearson's chi-squared ($\alpha = .05$)	Parametric bootstrap ($\alpha = .20$)
Strategy 4	LRT ($\alpha = .05$)	Parametric bootstrap ($\alpha = .20$)
Strategy 5	Pearson's chi-squared ($\alpha = .20$)	Parametric bootstrap ($\alpha = .05$)
Strategy 6	LRT ($\alpha = .20$)	Parametric bootstrap ($\alpha = .05$)
Strategy 7	Pearson's chi-squared ($\alpha = .20$)	Parametric bootstrap ($\alpha = .20$)
Strategy 8	LRT ($\alpha = .20$)	Parametric bootstrap ($\alpha = .20$)
Information criteria		
Strategy 9	AIC for all candidate models	
Strategy 10	BIC for all candidate models	

Model selection of GLMMs with SCED count data

Research questions

1) Which model selection strategies produce the least model selection bias?

2) Which model selection strategies yield the most accurate treatment effect estimates and reliable inferential statistics?

3)) Is there a relationship between model selection bias and the performance of the treatment effect estimator and inferential statistics?

Simulation Conditions

Parameter	Value
Series length (I)	10 (starting points of the intervention: 3, 4, 6, 7) or 20 (starting points of the intervention: 6, 8, 12, 14)
Number of cases (J)	4 or 8
Session length (T)	10 minutes
Baseline level (γ_{00})	log (0.05)
Treatment effect (γ_{10})	0, 1.79, 2.48, or 3
Between-case variance	
Baseline level (σ_{u0}^2)	1.0
Treatment effect (σ_{u1}^2)	1.0
Correlation (r_{u0u1})	−0.5
Poisson/NB model	
Log odds of excessive zeros (β_0)	$-\infty$
Treatment effect on excessive zeros (β_1)	$-\infty$
Dispersion parameter (θ)	2.0, 5.0 or $+\infty$
ZIP/ZINB model	
Log odds of excessive zeros (β_0)	−0.85, −1.39 or −2.19
Treatment effect on excessive zeros (β_1)	log (0.10)
Dispersion parameter (θ)	2.0, 5.0 or $+\infty$

A total of 16, 32, 48, and 96 conditions for data scenarios corresponding to the Poisson, NB, ZIP, and ZINB, respectively. In each condition, I simulated 2000 independent data sets (i.e., replications).

Data Analysis

Fitted Poisson, NB, ZIP, and ZINB models estimated by adaptive Gauss quadrature (AGQ) using the R package *GLMMadaptive*.

- The Wald test was adopted to conduct statistical inference for treatment effects.

Performance measures:

- Hit rate
- Bias, coverage rate of the treatment effect estimator
- Type I error rates

Performance Overview

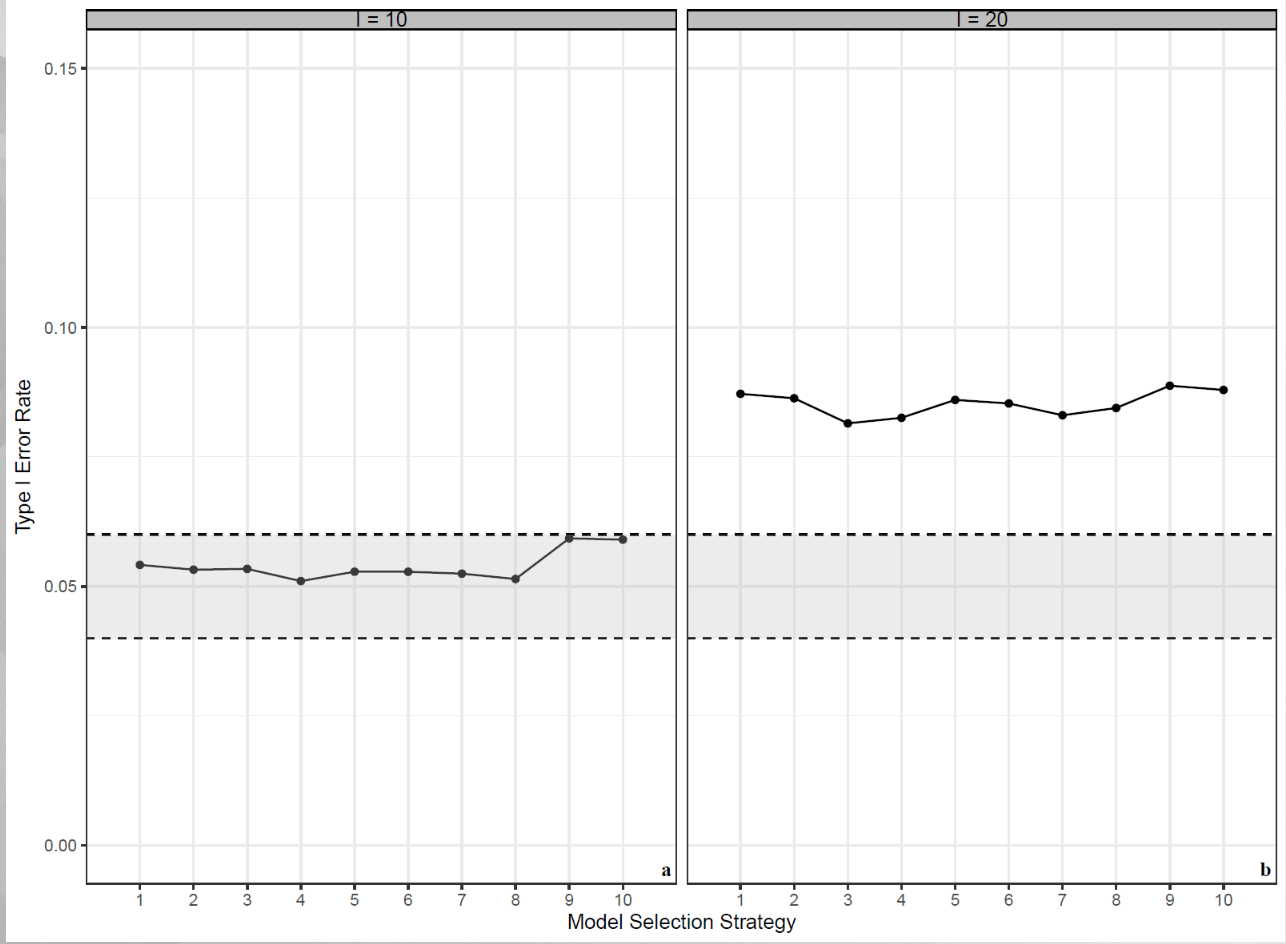
Strategy	Hit Rate (rank)	Bias (rank)	Coverage Rate (rank)	Type I Error Rate (rank)
Poisson and NB				
1. Pearson (α = .05) & PB (α = .05)	.777 (7)	.026 (8)	.913 (4)	.071 (8)
2. LRT (α = .05) & PB (α = .05)	**.854 (2)**	.027 (9)	**.914 (3)**	.070 (7)
3. Pearson (α = .05) & PB (α = .20)	.700 (10)	**.005 (1)**	.912 (8)	**.067 (2)**
4. LRT (α = .05) & PB (α = .20)	.775 (8)	**.007 (2)**	.912 (7)	**.067 (1)**
5. Pearson (α = .20) & PB (α = .05)	.825 (4)	.027 (10)	**.914 (2)**	.069 (6)
6. LRT (α = .20) & PB (α = .05)	**.865 (1)**	.026 (7)	**.914 (1)**	.069 (5)
7. Pearson (α = .20) & PB (α = .20)	.747 (9)	**.007 (3)**	.912 (6)	**.068 (3)**
8. LRT (α = .20) & PB (α = .20)	.790 (6)	.008 (4)	.912 (5)	.068 (4)
9. AIC	**.825 (3)**	.016 (6)	.911 (9)	.074 (10)
10. BIC	.816 (5)	.013 (5)	.910 (10)	.073 (9)

Performance Overview

Strategy	Hit Rate (rank)	Bias (rank)	Coverage Rate (rank)	Type I Error Rate (rank)
ZIP and ZINB				
1. Pearson ($\alpha = .05$) & PB ($\alpha = .05$)	.037 (7)	.249 (7)	.908 (10)	.081 (8)
2. LRT ($\alpha = .05$) & PB ($\alpha = .05$)	.031 (8)	.255 (9)	.909 (6)	.079 (6)
3. Pearson ($\alpha = .05$) & PB ($\alpha = .20$)	**.219 (3)**	**.156 (3)**	.910 (4)	**.073 (1)**
4. LRT ($\alpha = .05$) & PB ($\alpha = .20$)	.201 (4)	.172 (5)	**.910 (1)**	**.074 (2)**
5. Pearson ($\alpha = .20$) & PB ($\alpha = .05$)	.026 (9)	.253 (8)	.909 (7)	.080 (7)
6. LRT ($\alpha = .20$) & PB ($\alpha = .05$)	.025 (10)	.256 (10)	.909 (5)	.079 (5)
7. Pearson ($\alpha = .20$) & PB ($\alpha = .20$)	.195 (5)	.165 (4)	**.910 (3)**	**.074 (3)**
8. LRT ($\alpha = .20$) & PB ($\alpha = .20$)	.178 (6)	.177 (6)	**.910 (2)**	.076 (4)
9. AIC	**.232 (2)**	**.153 (2)**	.909 (8)	.082 (10)
10. BIC	**.254 (1)**	**.141 (1)**	.909 (9)	.081 (9)

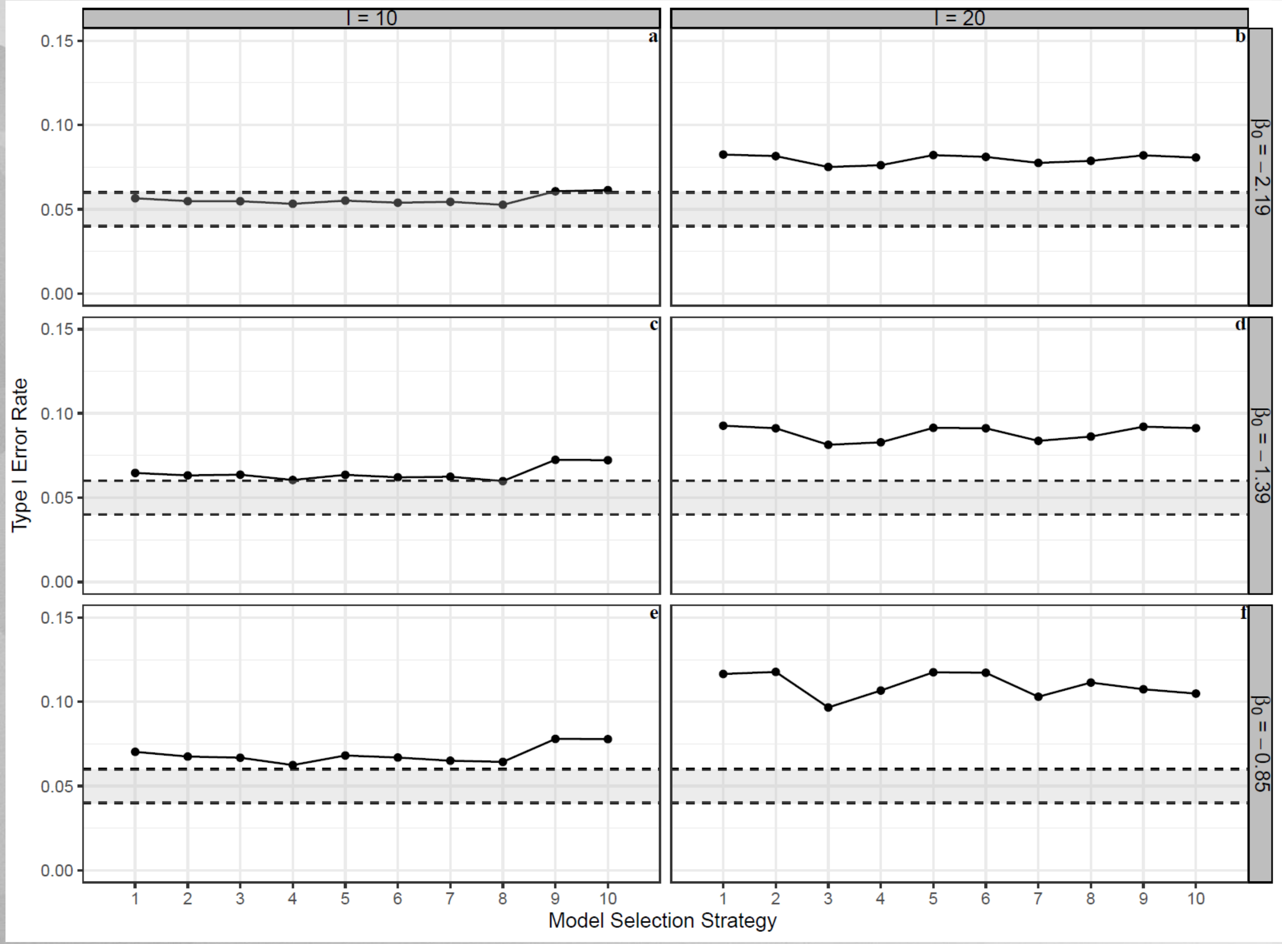
Type I Error Rate (Poisson and NB)

Type I Error Rate of Treatment Effects as a Function of the Model Selection Strategy and Series Length (I) for the Poisson and NB Data



Type I Error Rate (ZIP and ZINB)

Type I Error Rate of Treatment Effects as a Function of the Model Selection Strategy, Series Length (I) and Log Odds of Excessive Zeros (β0) for the ZIP and ZINB Data



Implications and Discussion

- When zero-inflation is not present, the absolute differences in performance measures among various model selection strategies were not practically significant.

- When zero-inflation is present, the overall low hit rates led to a very high proportion of incorrectly selected models, such as Poisson and NB models, which in turn caused unacceptable biased estimates of treatment effects for all strategies

Recommendations

When there are no zero observations in the outcome measurements:

- Using AIC and BIC to select an optimal model between NB and Poisson distributions to deal with SCED count data;
- If models are estimated by pseudolikelihood where AIC and BIC are not comparable, I recommend using Pearson's chi-squared test with a less penalty for model complexity (e.g., α=.20), as it has exhibited similar performance as AIC and BIC.

When zero observations are present in the outcome measurements:

- Using methods with a less penalty for model complexity (e.g., $\alpha = .20$) to accommodate zero-inflation. Information criteria, such as AIC and BIC, can also be adopted to compare all four candidate models.

Thank You