Examining the robustness of a design-comparable effect size in single-case designs

Qi-Wen Ding

Joint work with Li-Ting Chen, Yi-Kai Chen, Tong-Rong Yang, Yu-Shan Chiang, Cheng-Yu Hsieh, Che Cheng, Po-Ju Wu, Chao-Ying Joanne Peng

Multiple baseline design in SCED

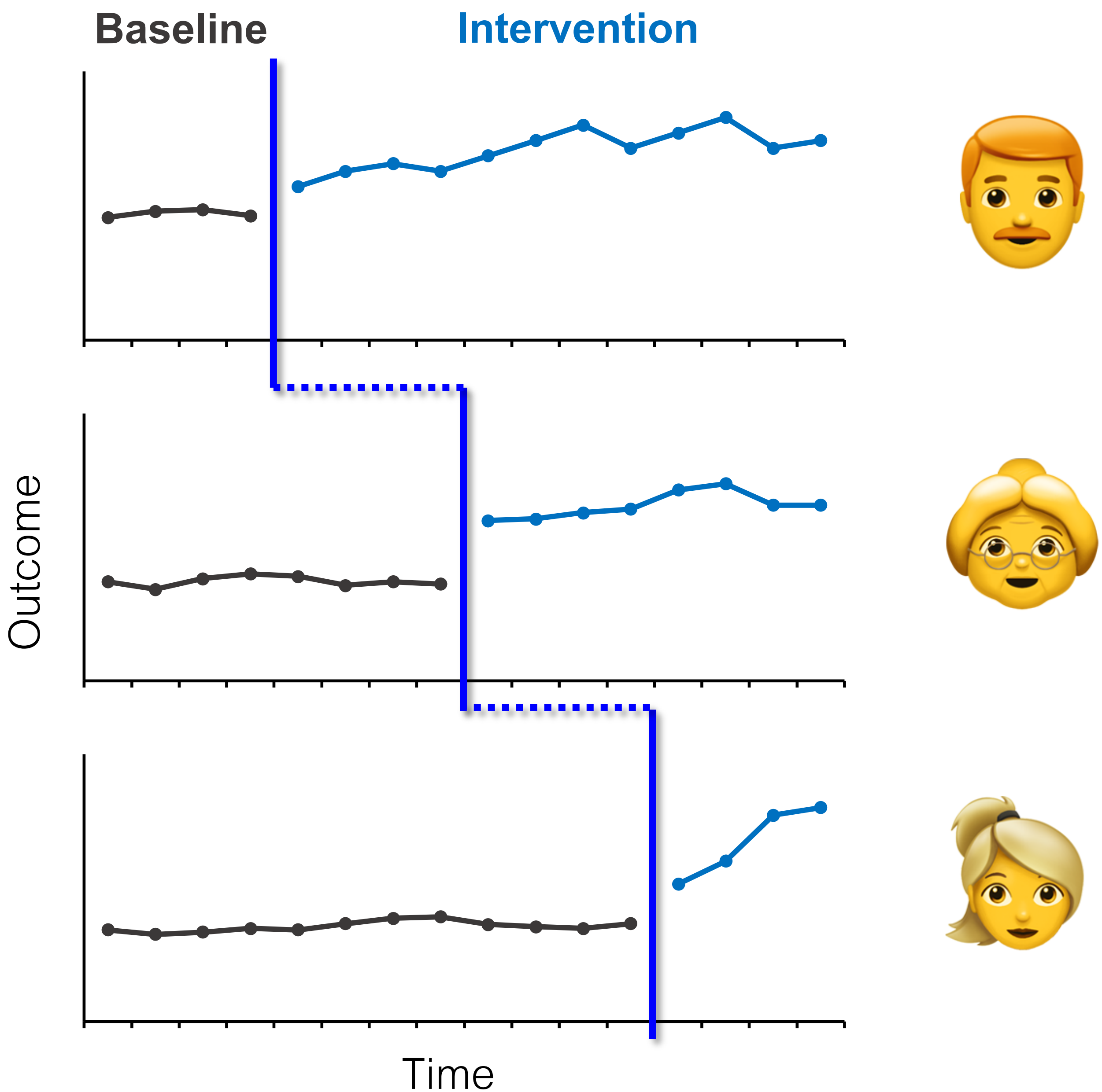
Single-case experimental design (SCED)

- Each participant serves as their own control.
- **Repeated measurements** over time.
- Small sample sizes.
- Focus on **within**-subject changes.

Multiple baseline (MB) design

- Most commonly-used type of SCED
- Two non-overlapping phases: phase A (baseline) and phase B (intervention)
- Each participant **begins the intervention at <u>different</u> time points**, which helps control for time-related effects
- Replicated across participants

MB design = intervention at staggered time points



Pustejovsky et al. (2014)'s measure

Linear mixed effect model: random intercept and slope

For the j^{th} measurement of the i^{th} participant, the outcome Y_{ij} is modeled by:

$$\begin{aligned} Y_{ij} &= \theta_{0i} + \theta_{1i} D_{ij} + \epsilon_{ij} \\ &= (\beta_0 + \gamma_{0i}) + (\beta_1 + \gamma_{1i}) D_{ij} + \epsilon_{ij}, \end{aligned}$$

where θ_{0i} = baseline level, θ_{1i} = individual binary treatment effect.
With below assumptions:

a) Within participant: $\boldsymbol{\epsilon}_i = (\epsilon_1, \ldots, \epsilon_j)^\intercal \sim MN(\mathbf{0}, \boldsymbol{\Sigma})$ where $\boldsymbol{\Sigma}$ follows AR(1) structure.

b) Across participants: $\mathrm{Var}(\epsilon_{ij}) = \mathrm{Var}(\epsilon_{hk}) = \sigma^2$ and $\mathrm{Cov}(\epsilon_{ij}, \epsilon_{hk}) = 0, \forall i \neq h$.

c) Random effects: $(\gamma_{0i}, \gamma_{1_i}) \sim MN(\mathbf{0}, \boldsymbol{T})$ with covariance $\boldsymbol{T} = \begin{bmatrix} \tau_0^2 & \tau_{10} \\ \tau 10 & \tau_1^2 \end{bmatrix}$.

d) Level-1 errors (ε_{ij}) are independent of random effects.

Design-comparable effect size under above model

$$\delta = \frac{\beta_1}{\sqrt{\sigma^2 + \tau_0^2}}$$, which is estimated by g_{AB} in eq(12) in our paper.

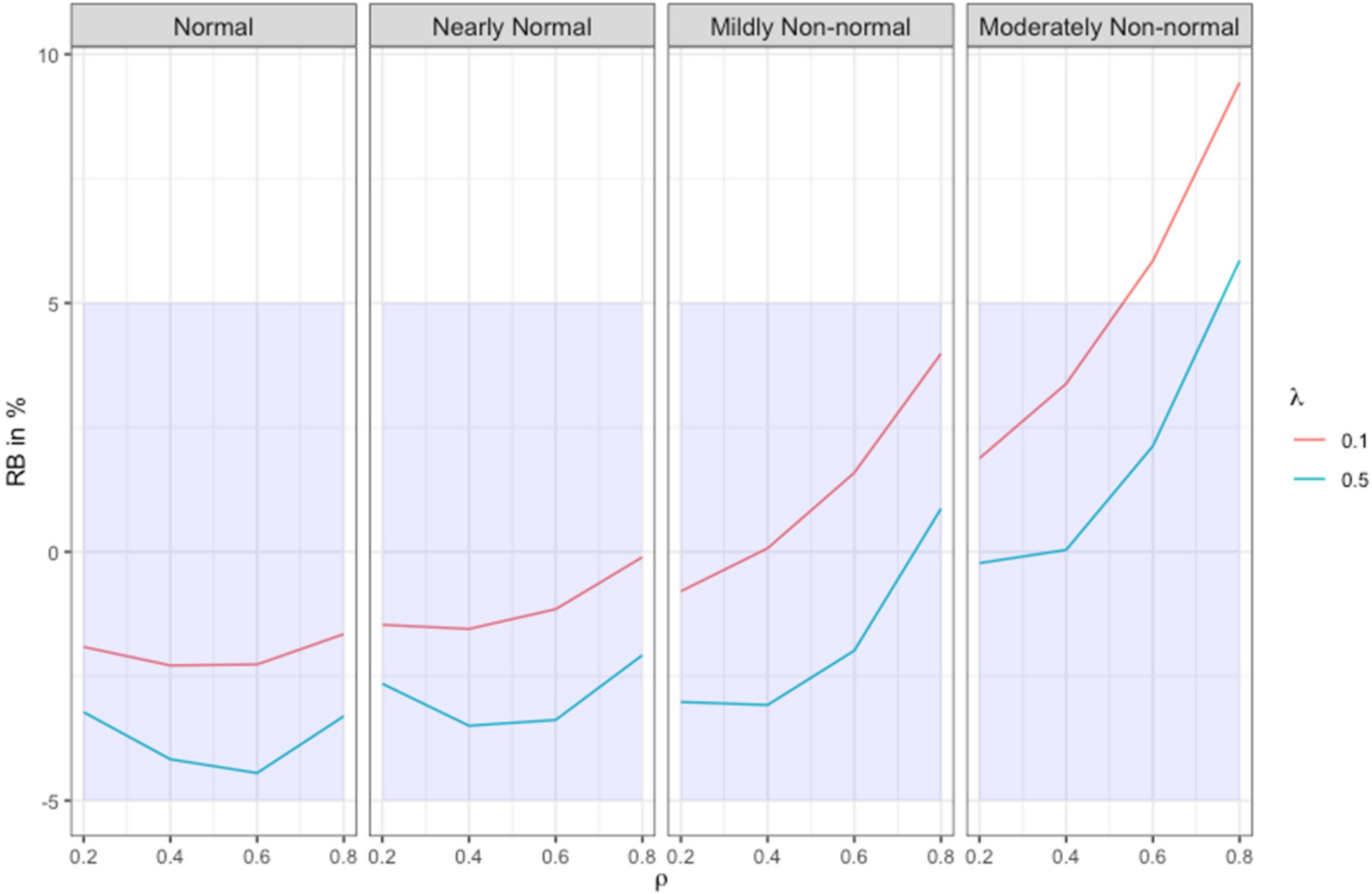
Simulation design

Factor	Definition	Conditions
Dist	Distribution of data	normal (0, 0)
	(skewness, kurtosis)	nearly normal (0, 0.35)
		mildly non-normal (1, 0.35)
		moderately non-normal (1, 3)
m	Number of participants	3, 4, 5, 6
N	Number of measurements	8, 16
ρ	Within-case reliability $= \tau_0^2/(\sigma^2 + \tau_0^2)$	0.2, 0.4, 0.6, 0.8
λ	Ratio of variance components $= \tau_1^2/\tau_0^2$	0.1, 0.5
ϕ	First-order autocorrelation	-0.4, -0.3, -0.1, 0, 0.1, 0.3, 0.4

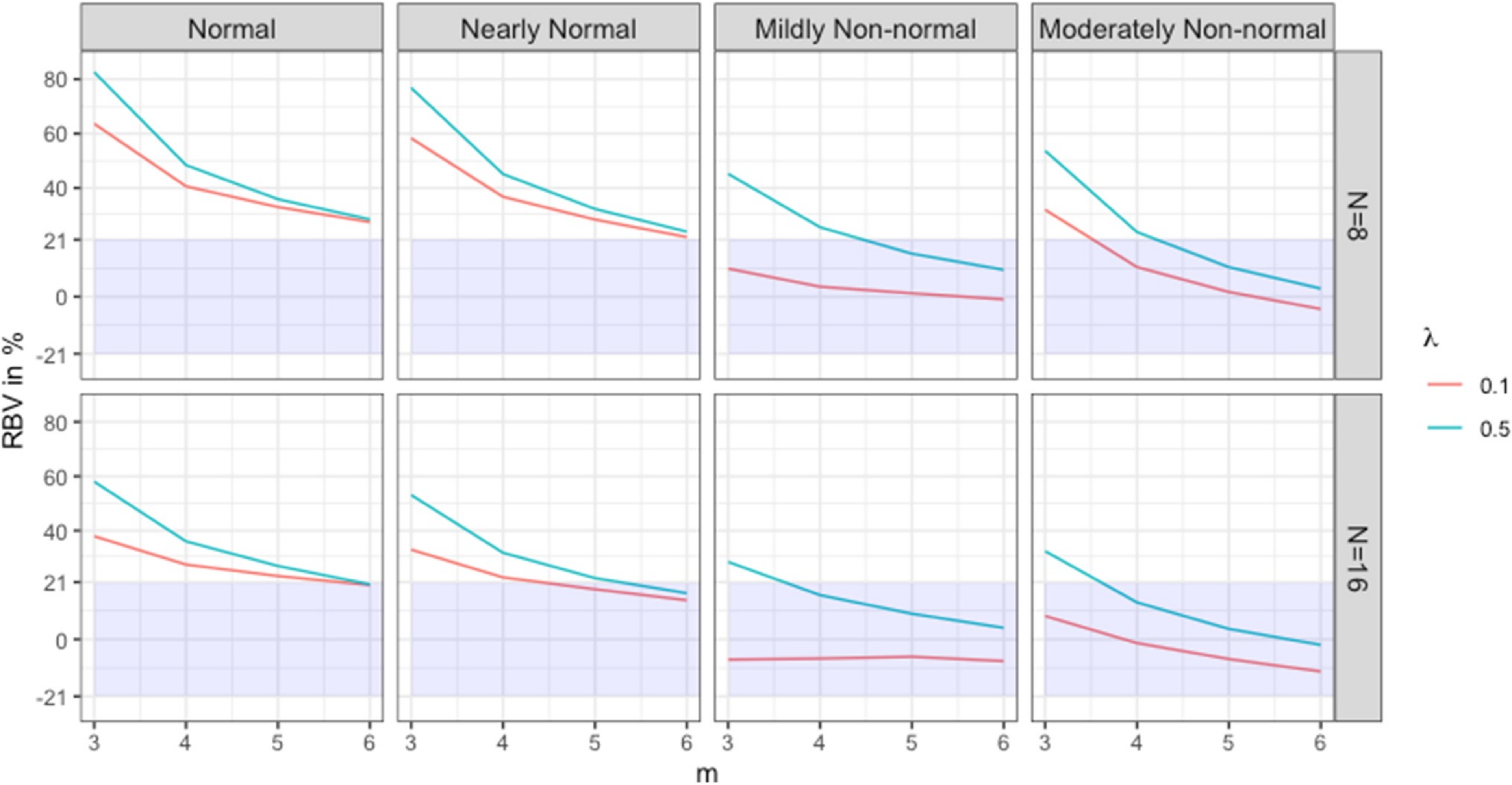
Evaluation criteria

- Relative bias of g_{AB}
- Relative bias of $\mathrm{Var}(g_{AB})$
- Mean square error of g_{AB}
- Coverage rate of symmetric C.I. for δ

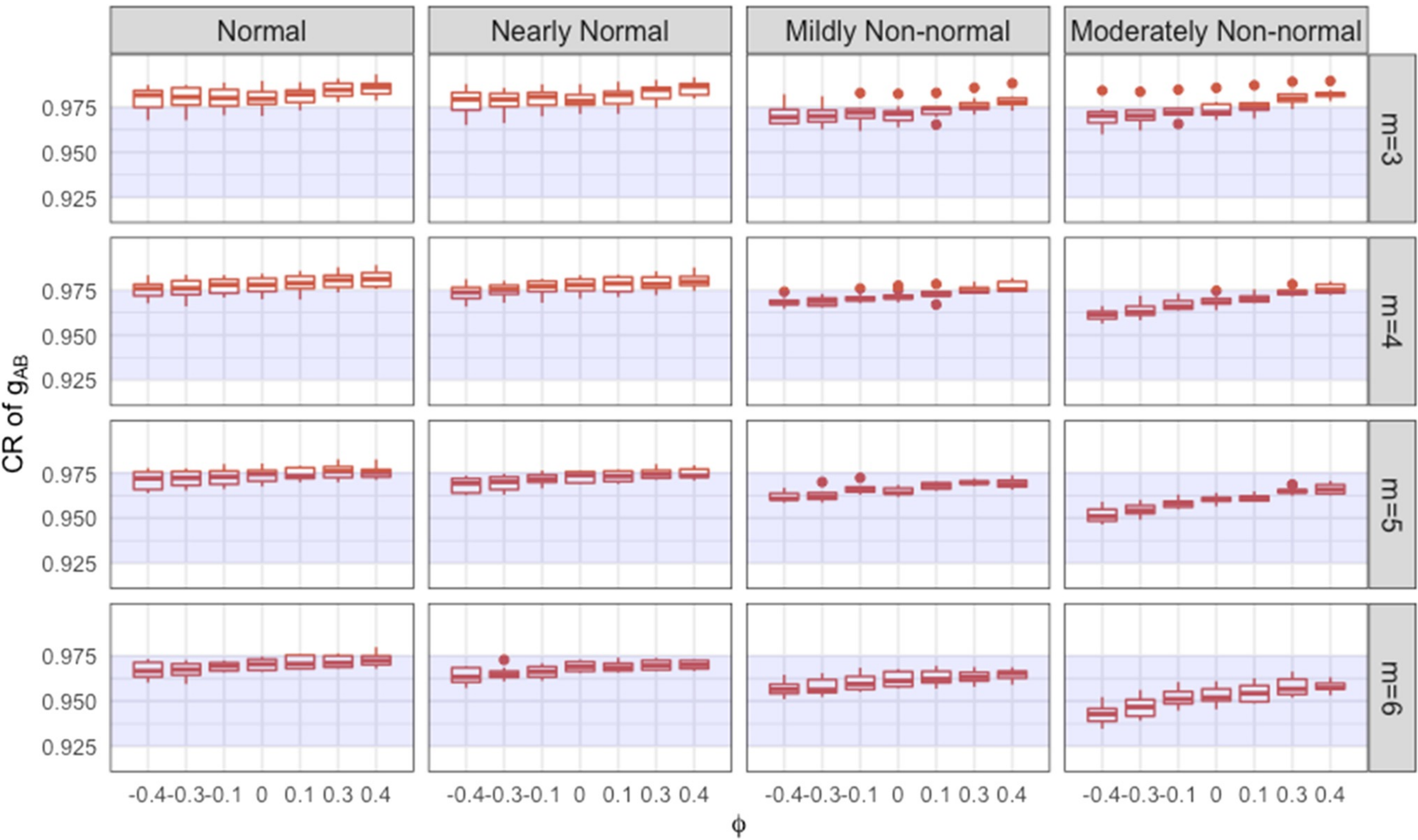
Results: relative bias of g_{AB}



Results: relative bias of Var(g_{AB})

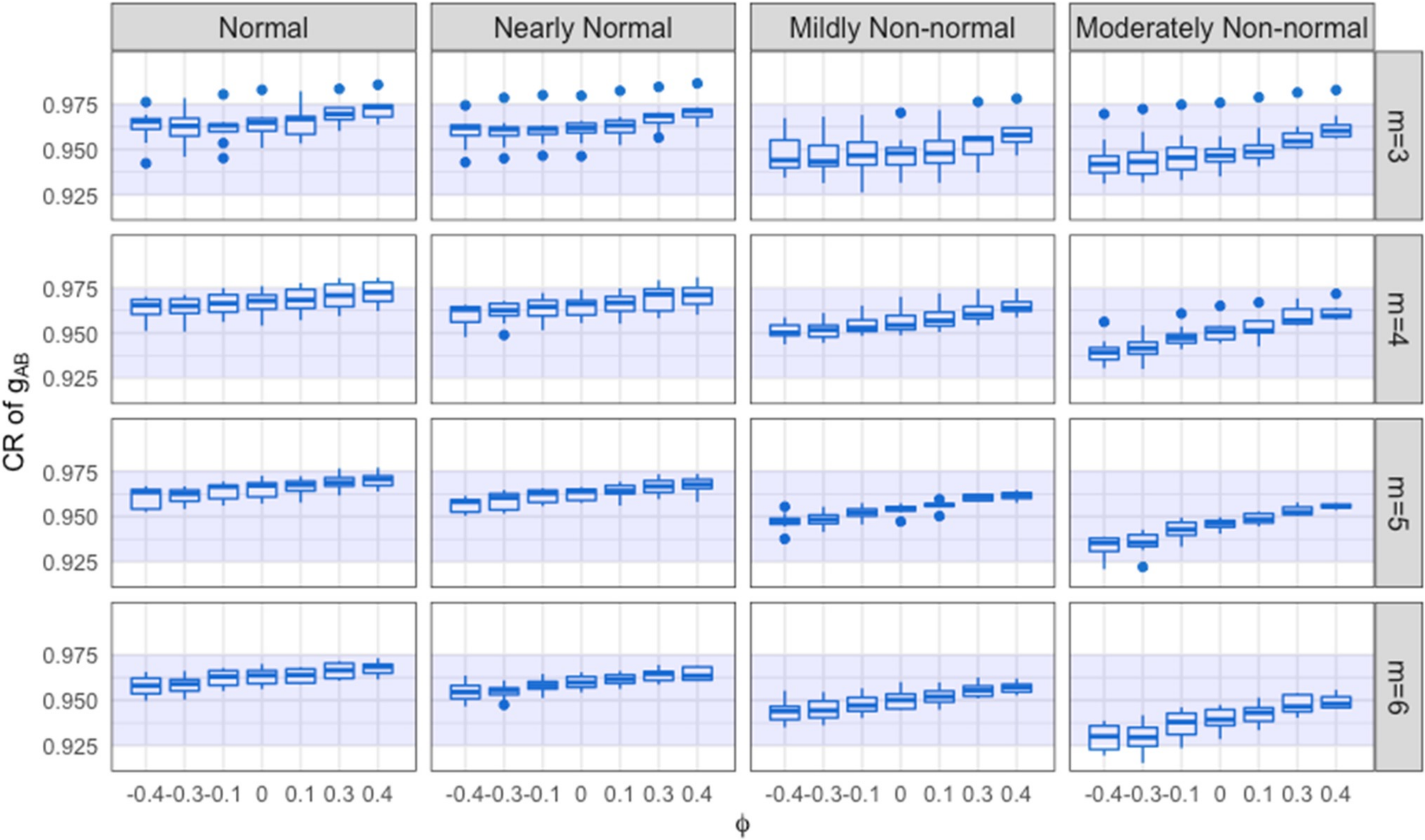


Results: coverage rate of C.I. ($N = 8$)



N: number of measurements m: number of participants

Results: coverage rate of C.I. ($N = 16$)



N: number of measurements m: number of participants

Summary

- Point estimation: unbiased even for non-normal cases
- Variance estimation: overestimated, especially for **normal** cases
- C.I.: over-coveraged, especially for **normal** cases

Recommendations

We recommend g_{AB} for MB studies and meta-analysis when each study with **at least 16 measurements**, meets one of the following conditions:

- $m = 6$, $\rho = .6$ or $.8$, and the shape, skewness, and kurtosis of data are similar to the normal or nearly normal
- $m \geq 4$, $\rho = .2$, $.4$, or $.6$, and the shape, skewness, and kurtosis of data are similar to the mildly or moderately non-normal distribution