Estimands and Analyses in Clinical Trials with Repeated Measures

Yan Wang, Ph.D.
Wonyul Lee, Ph.D.
Division of Biometrics IV
OB/OTS/CDER/FDA

***Views expressed here are those of the authors, not of the FDA**

S

The ASA Regulatory-Industry Statistics Workshop, September 22 – 25, 2020

Outline

- Hypothetical estimands addressed by mixed models for repeated measures
- Treatment policy estimands: slope-based estimands and fixed-time-point estimands

A Hypothetical Clinical Trial for Fabry Disease

Setting similar to many trials for slowly progressive rare diseases

- Randomized, double blind, active-controlled superiority trial
  - Control: received accelerated approval based on a biomarker endpoint
- Randomize (2:1) patients to receive biweekly intravenous infusion of **test product (TP)** or control for 12 months
- Patients with severe disease progression allowed to **switch treatment ("rescue")**
- Patients followed regardless of whether they discontinue their randomized treatment prematurely
- Efficacy outcome: estimated glomerular filtration rate (eGFR) measured at baseline and every 3 months
- **Primary endpoint:** change from baseline in eGFR **at 12 months**

Primary Efficacy Analysis: Mixed Model for Repeated Measures (MMRM)

- Fixed effects: treatment, **time (categorical variable),** interaction of time by treatment, and baseline value
- Unstructured covariance for within-subject correlations
- Not include the observed data after treatment discontinuation
  - Data after treatment discontinuation treated as missing
- Assume the data are missing at random (MAR)
  - The distribution of the unobserved data is well approximated by the distribution of the observed data
- Does not employ formal imputation for missing data

Supportive Analyses: all observed data are used regardless of treatment discontinuation

1. Same MMRM as used for the primary analysis
2. ANCOVA: with missing data imputed by the median value from the control group (Control-Median)

Efficacy Results: Mean Change from Baseline in eGFR at 12 Months

	TP LSmean (SE) (N=38)	Control LSmean (SE) (N=18)	Difference	
			95% CI	P-value
Primary Analysis (MMRM)	-1.7 (0.9)	-4.7 (1.2)	3.0 (0.1, 6.1)	0.04
Supportive Analyses*				
MMRM	-2.8 (1.0)	-4.9 (1.4)	2.2 (-1.2, 5.5)	0.21
ANCOVA (Control-Median)	-2.4 (0.9)	-4.9 (1.3)	2.6 (-0.7, 5.8)	0.12

* Include data after switch; Last observed value was used to impute data after death

Based on the primary analysis, one might conclude that the trial demonstrates a statistically significant treatment effect of TP on eGFR at 12 months

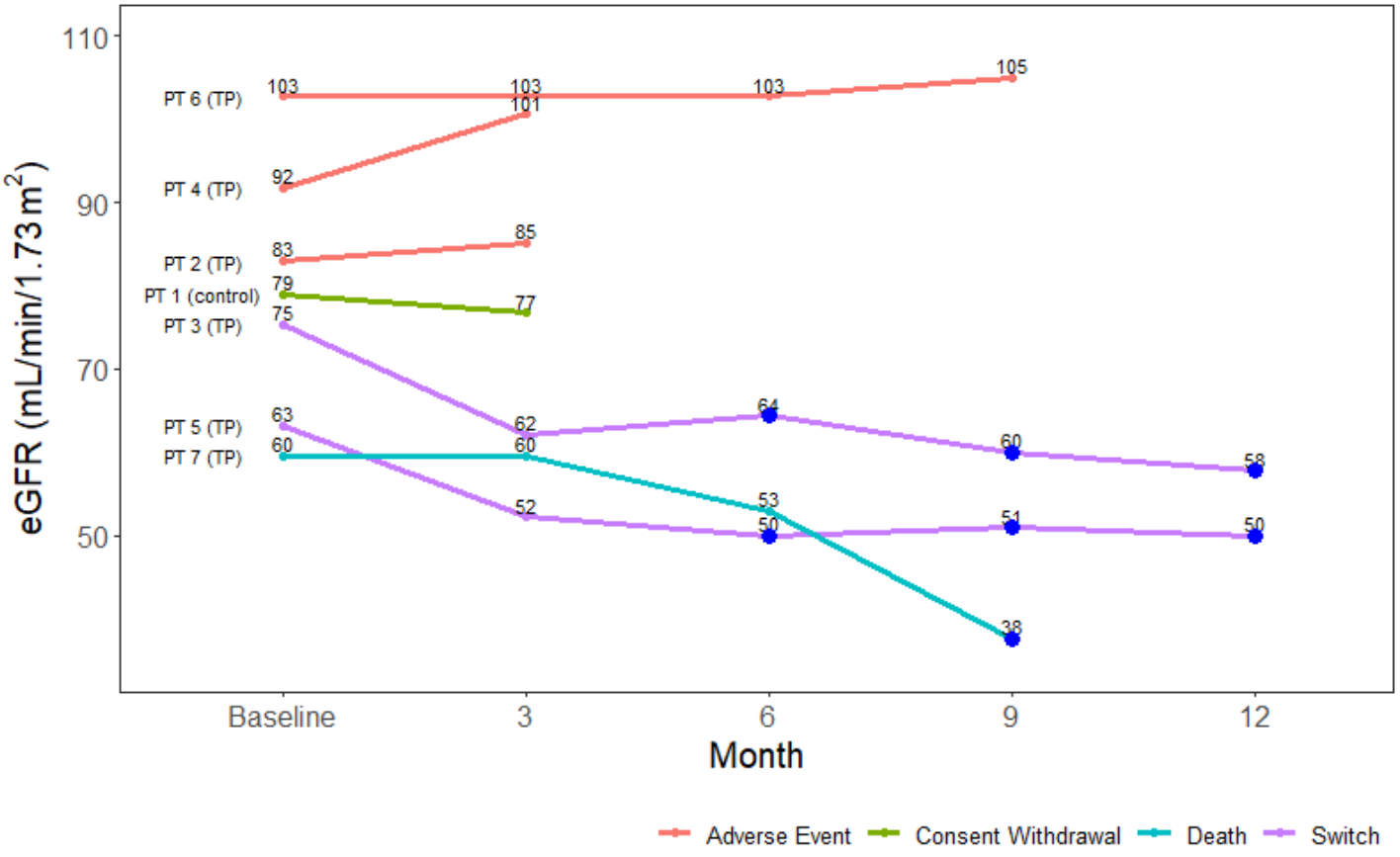
Treatment Effect (Estimand) Estimated by the Primary Analysis?

Hypothetical estimand: it concerns the treatment effect as if all patients had continued on randomized treatment for 12 months

- For patients who discontinue treatment, desired outcomes are the hypothetical values as if patients had not discontinued randomized treatment
  - These hypothetical values are treated as missing data in the MMRM analysis

Intercurrent Events: Treatment Discontinuations

	Lack of Efficacy	Adverse Events	Consent Withdrawal
TP	3 (one died after treatment switch)	3	0
Control	0	0	1



Blue dots indicate values after treatment switch

MMRM (Hypothetical Estimand): Implicitly Impute Values for Patients Who Discontinue TP

Reason of discontinuation	Lack of Efficacy			Adverse Events		
PT #	PT 3	PT 5	PT 7	PT 2	PT 4	PT 6
Last observed Δ in eGFR	-17	-13	-22	2	9	2
Imputed Δ in eGFR at 12 m	-6	-7	-9	0	1	1

Would have much better outcomes if patients had continued TP

Such hypothetical scenarios are not clinically plausible because these patients discontinue TP due to lack of efficacy

The hypothetical treatment effect (estimand) estimated by the MMRM is not recommended

Supportive Analyses: use all observed data regardless of treatment discontinuation → estimate treatment policy estimand

Efficacy Results

Supportive Analysis	TP LSmean (SE) (N=38)	Control LSmean (SE) (N=18)	Difference	
			95% CI	P-value
MMRM	-2.8 (1.0)	-4.9 (1.4)	2.2 (-1.2, 5.5)	0.21
ANCOVA (Control-Median)	-2.4 (0.9)	-4.9 (1.3)	2.6 (-0.7, 5.8)	0.12

Discontinuation due to Adverse Events			
PT #	PT 2	PT 4	PT 6
Last observed Δ in eGFR	2	9	2
MMRM: Imputed Δ in eGFR at 12 m	1	1	-1
Control-Median Imputed Δ in eGFR at 12 m	-2	-2	-2

Using MMRM to estimate the treatment policy estimand may be less concerning

Summary: Estimands Estimated by MMRM

- MMRM is one of the most commonly proposed primary analysis methods, targeting either a hypothetical estimand or the treatment policy estimand
- Major concern with using MMRM to estimate hypothetical estimands
  - No rationale is provided on why MMRM is a reasonable approach for handling "hypothetical values" for patients who discontinue study treatment due to lack of efficacy or adverse events
- Using MMRM to estimate treatment policy estimand may not be concerning if the amount of missing data is small

Two Types of Treatment Policy Estimands Clinical Trials for Fabry Disease

- **Fixed-time-point estimand:** Treatment difference in the **mean change from baseline in eGFR at a pre-defined fixed time point**
- **Slope-based estimand:** Treatment difference in the **mean annualized rate of change in eGFR (slope)**
  - Assume that the trajectory of eGFR is approximately linear
  - Reasonable assumption for slowly progressive rare diseases

Question: When the outcome trajectory is linear, should one select the **slope-based estimand** as the primary estimand?

Analysis Approaches for Slope-based Estimand

One-stage vs. Two-stage

- One-stage approach: repeated measures are analyzed using a linear model that includes **time as continuous variable**
  - When the model **includes** intercept and slope as random effects, within-subject residual errors are typically assumed to be independent
  - When the model **does not include** random effects, within-subject correlations are characterized by a covariance structure, such as compound symmetric, autoregressive, or unstructured covariance

- Two-stage approach:
  - ***Stage 1:*** A slope is estimated for each subject from the simple linear regression of the subject's outcome variable on time
  - ***Stage 2:*** ANCOVA (or non-parametric methods) is used to estimate the mean slopes between groups

Slope-based Estimand
Which analysis approach should be used?

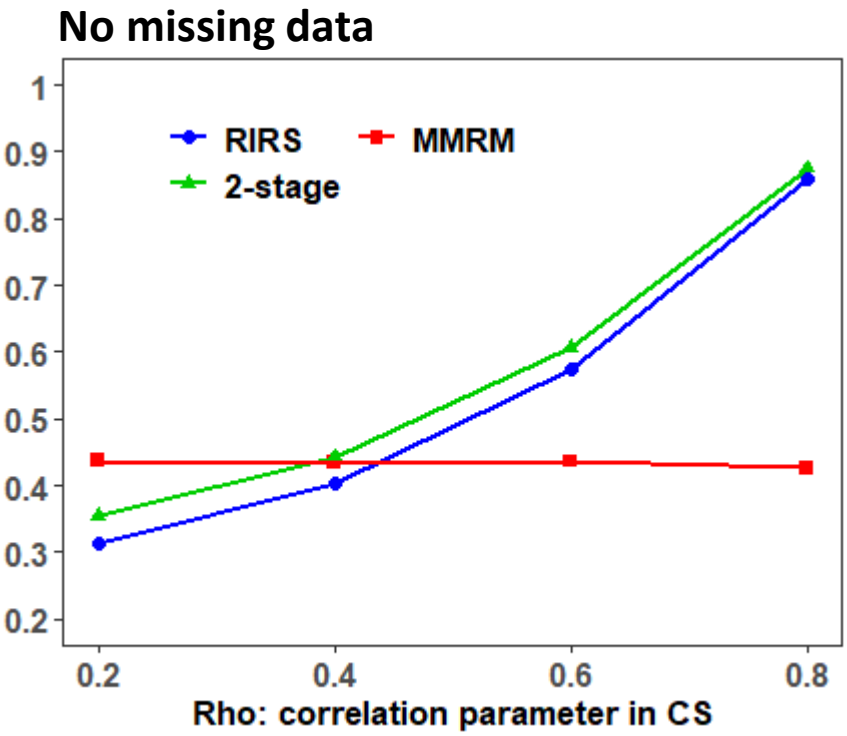
- One-stage approach:
  - **Higher power to detect a treatment difference compared to two-stage?**
- Two-stage approach:
  - **No assumption on within-subject correlations!**
  - **Lower power compared to one-stage approach?**

Simulation Study Evaluating Performance (Power) of Slope-based and Fixed-time-point-based Approaches

- Slope-based estimand:
  - **One-stage**: a linear model with random intercept and random slope → **RIRS**
  - **Two-stage**
- Fixed-time-point estimand: MMRM (**time is treated as categorical variable**) and ANCOVA
- Small sample size: Test Product (TP) = 30 & Control = 15
- Replications: 5000

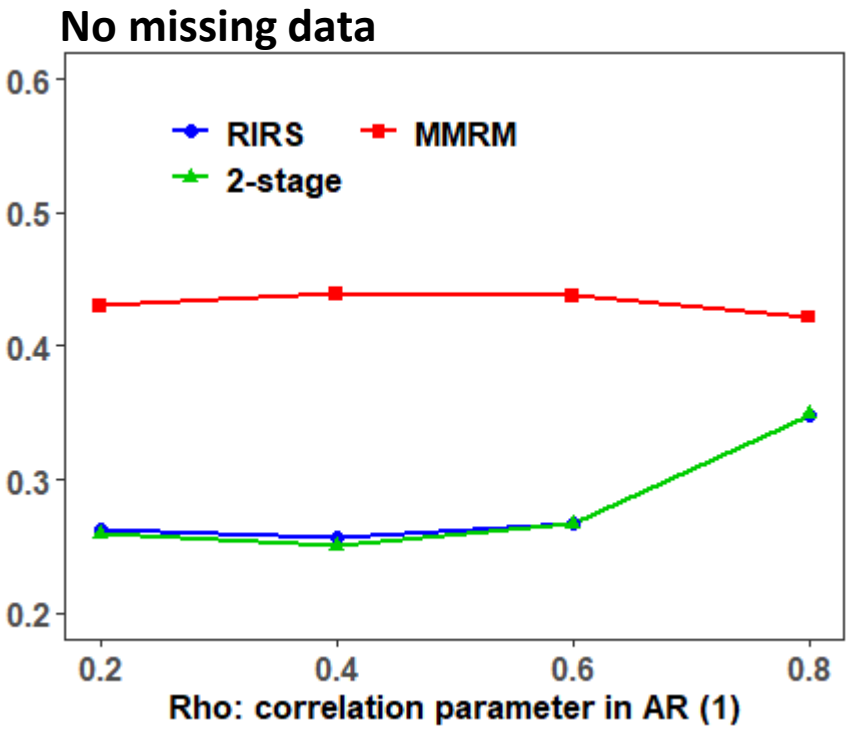
Case 1: Data are generated under a linear model with a covariance of compound symmetry (CS)

- Slope-based methods: **power ↑ as correlation ↑**
- **2-stage method has higher power than 1-stage method (RIRS)** when correlation < 0.8
- Fixed-time-point method **have higher power** than **slope-based methods** when correlation is small (0.2)



Case 2: Data are generated under a linear model with a covariance of autoregressive 1 (AR (1))

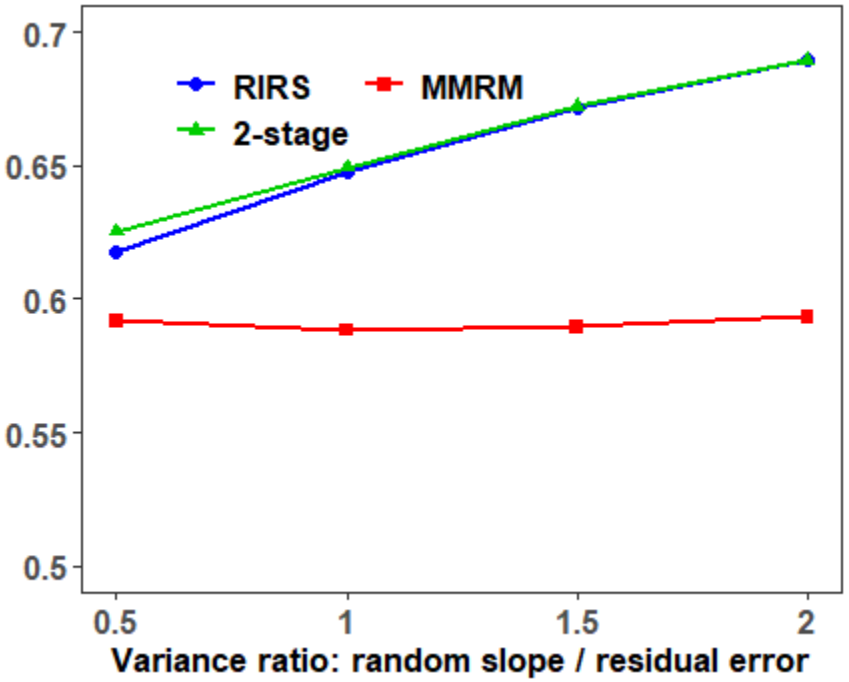
- Slope-based methods: **similar power; power ↑ as correlation ↑**
- Fixed-time-point method **have higher power** than **slope-based methods**



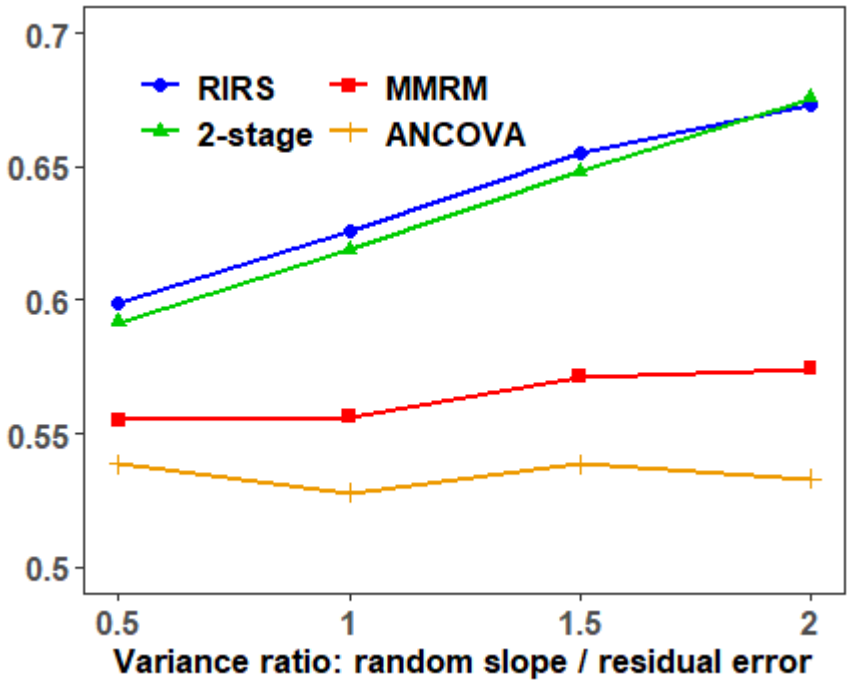
Case 3: Data are generated under RIRS

Var (random intercept) = **0.5** ∗{ Var(random slope) + Var(residual error) }

No missing data



Missing data (MAR): 10%

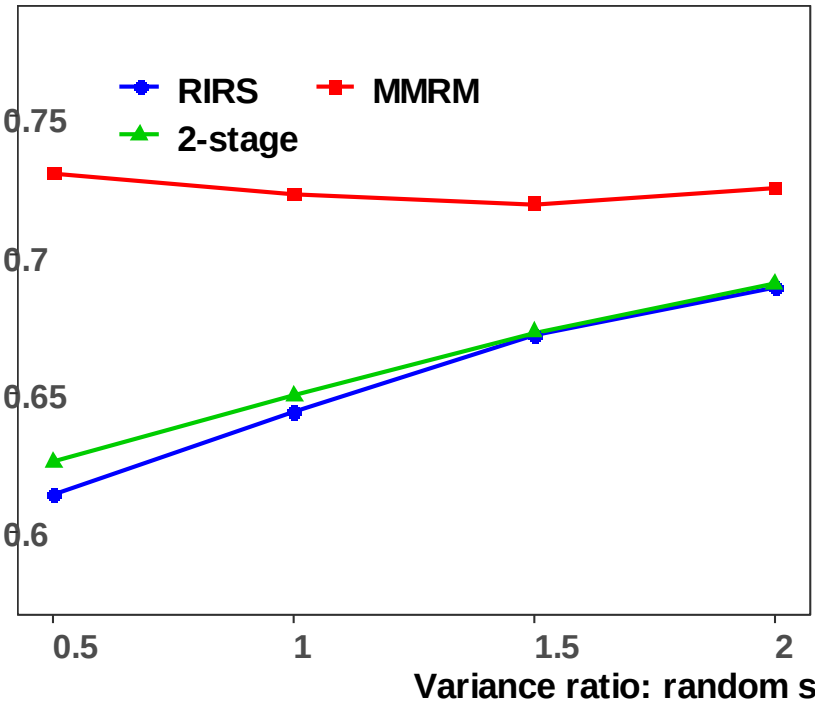


- **Slope-base methods outperform** fixed-time-point method
- Slope-based methods: **power ↑ as variance ratio ↑**
- Without missing data, **2-stage performs as good as RIRS**
- With missing data, RIRS performs slightly better when variance ratio ≤ 1.5

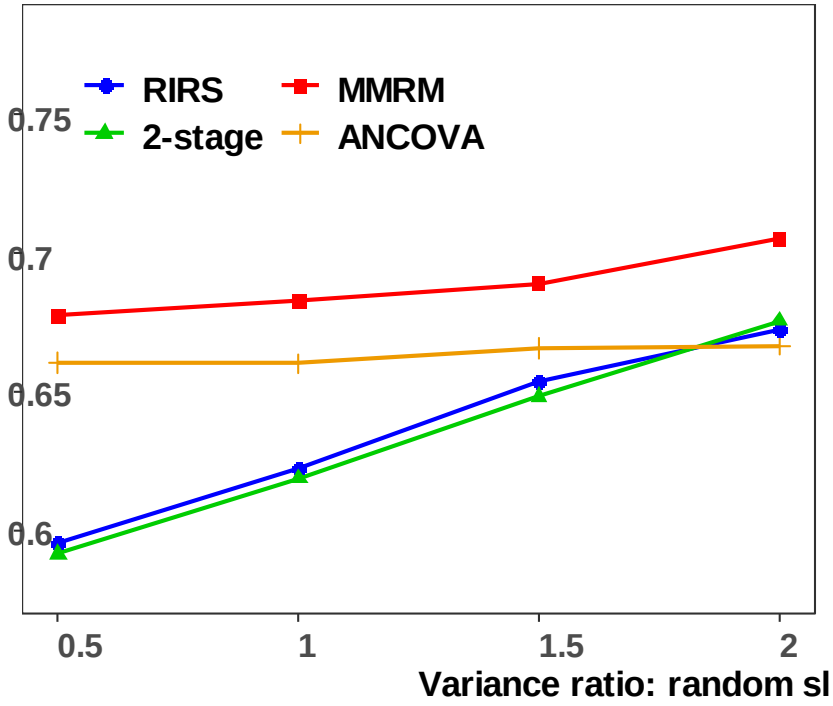
Case 4: Data are generated under RIRS

Var (random intercept) = **0.1** ∗{ Var(random slope) + Var(residual error) }

No missing data



Missing data (MAR): 10%

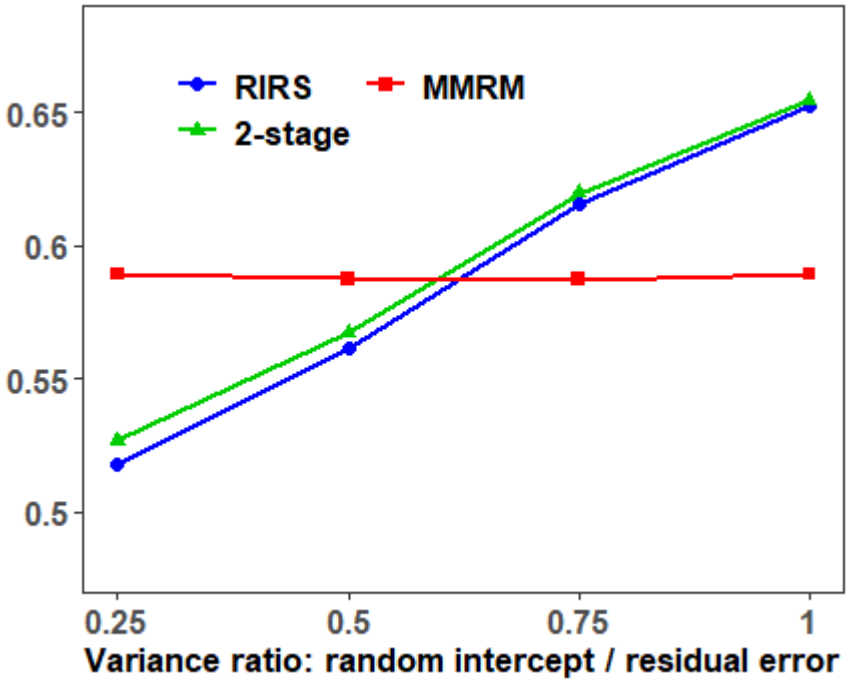


- Fixed-time-point method (MMRM) **outperforms slope-base methods**
- Slope-based methods: **power ↑ as variance ratio ↑**
- **2-stage performs as good as one-stage (RIRS)**

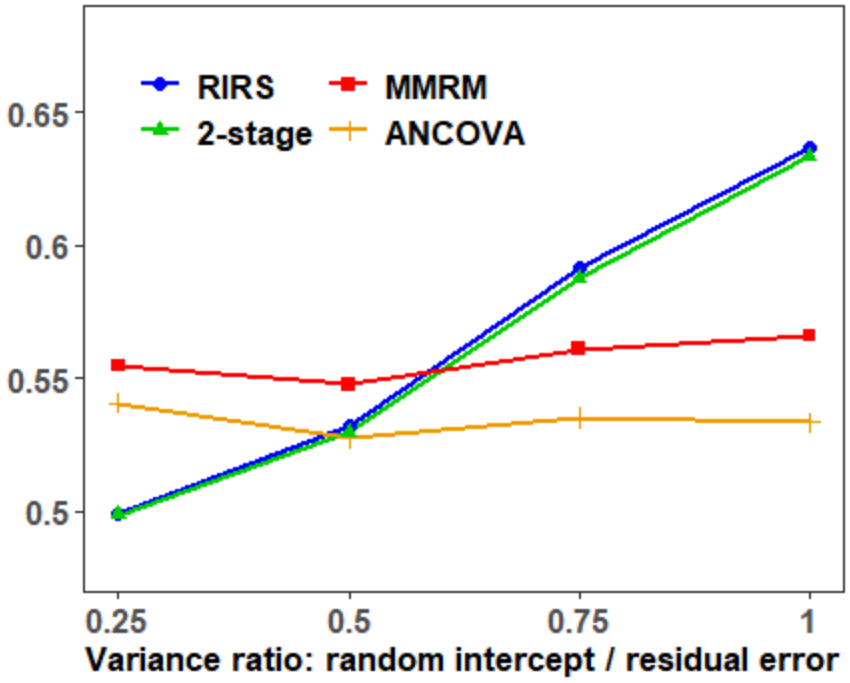
Case 5: Data are generated under RIRS

Var (random slope) = **0.5** ∗{ Var(random intercept) + Var(residual error) }

No missing data



Missing data (MAR): 10%

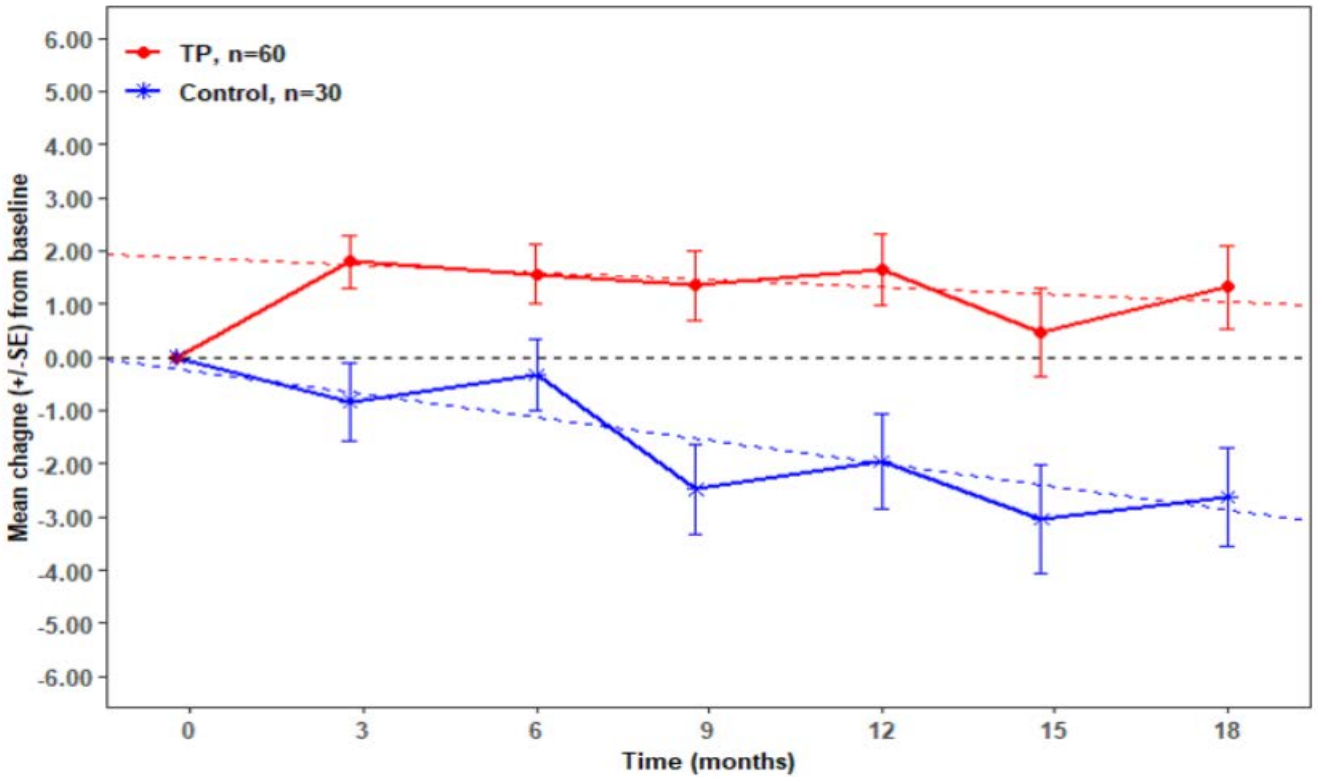


- Slope-based vs. fixed-time-point: **no consistent winner**
- RIRS vs. 2-stage: **performs similarly**

Application: Hypothetical Clinical Trial for Fabry Disease

Trial objective is to compare the mean slopes of eGFR change from baseline between two treatment groups

Mean Change in eGFR



~8% missing data at 18 months

Application: Hypothetical Clinical Trial for Fabry Disease

Efficacy Results

	TP LSmean (SE) (N=60)	Control LSmean (SE) (N=30)	Difference	
			95% CI	P-value
Slope-based Approach				
Primary: RIRS (Kenward-Rogers)	-0.5 (0.5)	-1.7 (0.7)	**1.2 (-0.6, 3.0)**	**0.187**
Sensitivity: two-Stage	-0.6 (0.6)	-2.2 (0.8)	**1.6 (-0.4, 3.6)**	**0.106**
Fixed Time Point Approach (Post Hoc)				
MMRM	1.3 (0.7)	-2.3 (1.0)	**3.7 (1.2, 6.1)**	**0.004**
ANCOVA	1.3 (0.7)	-2.7 (1.2)	**4.0 (1.3, 6.7)**	**0.004**

- Slope-based methods: RIRS yields **less favorable results** compared to 2-stage method
- Slope-based methods yield **much less favorable results** compared to fixed-time-point methods

Additional Simulation: Based on Application of RIRS

Sample Size: TP=60 vs. Control =30; **Estimated slope:** TP=-0.5 vs. Control=-1.7

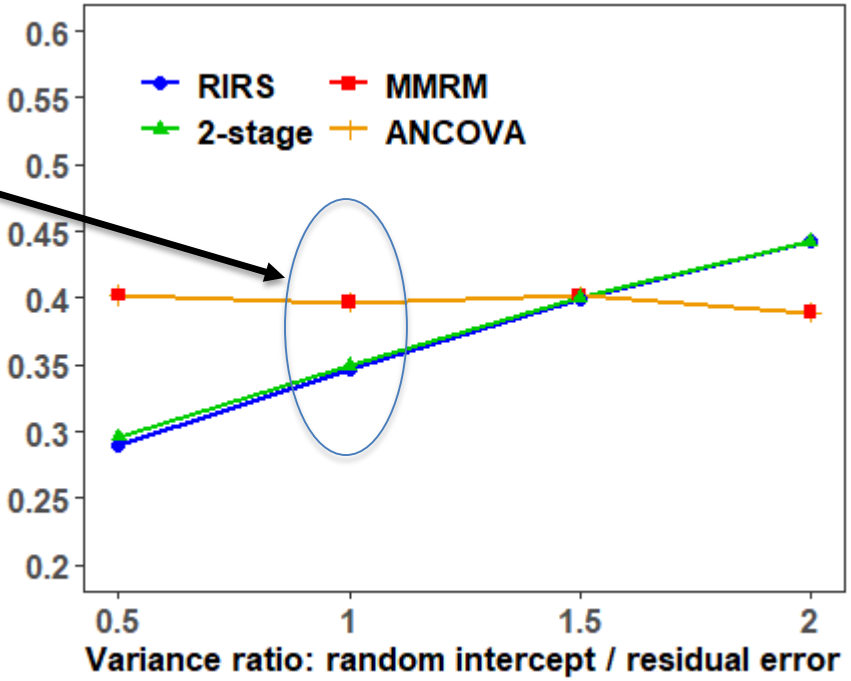
Estimated variances:

Random intercept = 7.7, Random slope = 7.0, Residual error= 7.9

Variance ratio = **Var(random intercept)** / Var(residual error) = **0.98**

Var (random slope) = **0.45** ∗{ Var(random intercept) + Var(residual error) }

Fixed-time-point methods
outperform slope-based methods



Conclusions

- In terms of power, the performance of slope-based analyses depends on the underlying covariance structure of the repeated measures → **power ↑ as correlation ↑**
- Slope-based: One-stage vs. two-stage:
  - Without missing data, two-stage approach performs as good as (if not better) one-stage approach
  - With missing data (MAR), one-stage approach slightly performs better than two-stage approach
- Slope-based vs. fixed-time-point: no approach universally outperforms the other