Navigating Complexity: Joint Modeling of Multiple Longitudinal Endpoints and Simulation Guided Adaptive Designs in Clinical Trials

Liangcai Zhang, Statistical Modeling & Methodology
Statistics & Decision Sciences, Quantitative Sciences of Clinical Pharmacology

2025 Annual ASA NJ Chapter Spring Symposium

June 13, 2025

Johnson&Johnson
Innovative Medicine

Acknowledgement

Grateful Recognition to:

Organizers
American Statistical Association (ASA) NJ Spring Chapter Committee:
For organizing this exceptional event and providing a platform to share our work

Collaborators
J&J Innovative Medicine – Matrix Team (Statistics, Clinical, Modeling, Quantitative Sciences and Biomarker etc) Across Therapeutic Areas
For their valuable input, collaboration, and support in advancing this research

Special Thanks
To **my team and co-authors** for their dedicated efforts and impactful contributions
George Capuano, Vladimir Dragalin, John Jezorwski, Kim Hung Lo and Fei Chen

To **symposium attendees and fellow statisticians** for the engagement and inspiring discussions

Johnson&Johnson
Innovative Medicine

Table of contents

Introduction

Modern clinical trials demand precision, adaptability and patient-centric approaches that integrate increasingly complex data sources

- ❑ Patient-centric: Better insights into treatment effects across multiple dimensions + over time courses + enhanced patient safety via continuously monitoring

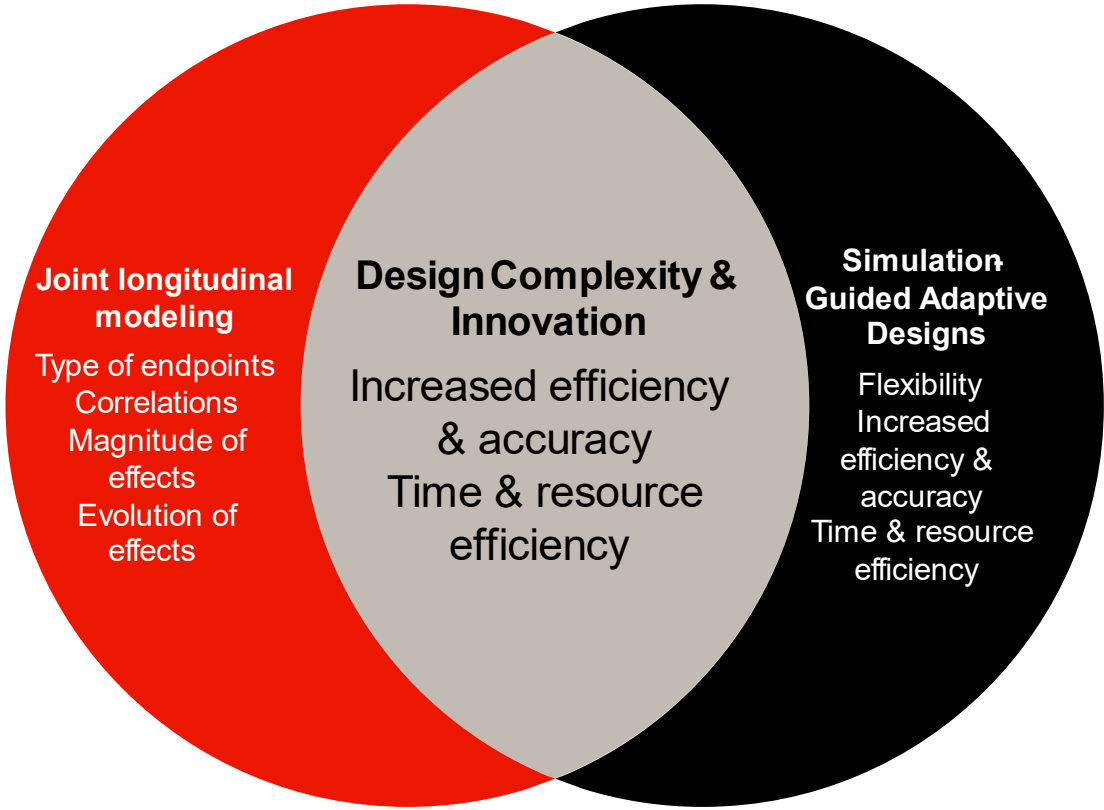
  Joint modeling + longitudinal modeling

- ❑ Efficiency (rapid insights): Improved resource utilization and faster trial completion times

  Design Adaptations

- ❑ Risk management: powerful models to design and predict outcomes based on hypothetical scenarios

  Simulation-guided



allowing real-time, data-driven adjustments ensuring trial integrity, e.g., adaption of adding more patients in promising treatment arms per early results

Simulation-Guided Adaptive Clinical Trials – planning a trial

representing powerful & modern methodology designed to improve trial efficiency, flexibility, and decision-making by leveraging advanced computational simulations

Adaptive clinical trials allows modifications to the study design as trials progress based on accumulating data, without compromising statistical validity or integrity

- ❑ Adjusting sample size
- ❑ Modifying treatment arms
- ❑ Dropping ineffective interventions

Simulation-guided approach would help

- ❑ Properly understand the design and get a good understanding how your trial may perform, e.g., the probability of success over a range of possible 'truth'
- ❑ Remove the constrains of closed form analysis
- ❑ Clearly communicate more sophisticated methods at planning stage to support the cross-functional team

Joint Modelling of Multiple Longitudinal Endpoints – monitoring a trial

enabling more comprehensive understanding of treatment effects and disease progression & analyzing complex relationships between endpoints measured repeatedly over time

Joint modeling simultaneously accounts for the interdependencies between endpoints, and offers holistic understanding of treatment effects

- ❑ Maximizes data utilization across endpoints and avoids redundant analysis – improved efficiency
- ❑ Captures nuanced relationships for more patient-centric decision-making – precision insights
- ❑ Provides better early predictions of clinical outcomes – enhanced predictive power
- ❑ Enables comprehensive evidence to satisfy regulatory demands – regulatory confidence

For example, multiple longitudinal endpoints (e.g., tumor size, quality of life) and their dependency on clinical outcomes

Motivating example for monitoring/analyzing a trial

Dose-response analysis

In clinical development, dose response is an important question to be addressed when longitudinal data are collected. Dose response analysis could be characterized by

- Primary endpoint (or key efficacy endpoints) at marginal level
- Comprehensive composite endpoint profiles at joint level

Suppose a phase 2 trial collects 4 longitudinal endpoints:

Endpoint Category	**Endpoint Name**	**Data Type**
Primary	Endpoint A	Continuous
Key Secondary	Endpoint B	Continuous
	Endpoint C	Binary
Mechanism-of-Action (MoA)	Endpoint D	Continuous

Schedule of activity (SoA): every 2 weeks up to week 24; Multiple doses: high dose, low dose and placebo/control

Motivating example for monitoring/analyzing a trial

Dose-response analysis

Questions of interest:

- How the variation of an endpoint (e.g. biomarker, MoA) is related to the variation of another – correlation over time between endpoints
- How to model the association evolving over time – autocorrelation within endpoints
- How to best utilize all available data to inform dose(s) – joint and/or longitudinal modeling vs marginal modeling

Motivating example for planning a trial

Design adaptation and power boosting

For a proof-of-concept phase 2 trial, adaptive feature of having interims plus flexibility of claiming success at IA would be recommended by leveraging rich longitudinal data due to factors like the competitive development landscape

- How it compares with traditional method?
- Concerns on false positive inflations?

Questions of interest:

- How the power can be boosted at earlier time (say planned IA) – power gain by leveraging data over time and across endpoints
- Potential sample size saving and/or shorter trial duration – possibilities of early efficacy claiming

Method

"Totality of information" –> precision of estimates and increase of study power & efficiency

- Shared term strategy & conditional probability model

$$Y_1 \sim g(Z, Visit, \ldots)$$

$$Y_2 \mid Y_1 \sim f(Y_1, Z, Visit, \ldots)$$

- Marginal processes & joining in one way or another via latent variable *etc*

$$Y_1 \sim a_1 + b_1(Zi) * Visit(j)$$

$$Y_2 \sim a_2 + b_2(Zi) * Visit(j)$$

$$(a1, a2) \sim BVN(., G_a(S_a, rho_a)) \text{ AND } (b1, b2) \sim BVN(., G_b(S_b, rho_b))$$

Statistical models

- Endpoints are modeled by Emax model (flexible for binary/cont.)
  - Y_{ijk} - endpoint *k* in patient *i* at visit *j*
  - $Y_{ijk} \sim N(e_0[z_i,k] + emax[z_i,k] \cdot \frac{visit_j}{visit_j + et50[z_i,k]}, \mathrm{V}^{\mathrm{k}}_{J\times J})$; *Trunc*(L, U)
  - $\mathrm{V}^{k}_{J\times J}$ has *AR(1)* auto-correlation structure, (or compound symmetry, or general)

- Priors
  - $e_0[z_i, 1{:}K] \sim N(rep(0,K), \Omega^1_{K\times K})$
  - $emax[z_i, 1{:}K] \sim N(rep(0,K), \Omega^2_{K\times K})$
  - $et50[z_i, 1{:}K] \sim N(rep(5,K), \Omega^3_{K\times K})$

- Priors on variability
  - $\Omega^1_{K\times K} \sim Wishart(diag(K), K)$
  - $\Omega^2_{K\times K} \sim Wishart(diag(K), K)$
  - $\Omega^3_{K\times K} \sim Wishart(diag(K), K)$

Correlated Emax parameters through multi-variate normal distributions

Note: (1) We are not modeling/fitting dose (x) and response (y) directly, but more on 'better' estimating over-time treatment effects (2) Flexible inputs of endpoint types: binary, continuous, even easy to extend to time-to-event using zeros-tricks

Application to Planning Trial Design:

Bivariate model of functional and structural endpoints in Ph2 PoC study

Real-world data indicates functional endpoint (primary) is strongly linearly related to structural endpoint (secondary)

Strategic value of Functional endpoints

- **Regulatory agencies:** emphasize to sponsors that the preferred endpoints link treatment benefit to how a patient feels or functions in disease-related activities
- **Differentiation:** Competitor trials to date have not shown a benefit in functional activities
- **Our strategy:** Potentially allow differentiation between assets

Strategic value of Structural endpoints

- **Regulatory agencies:** Primary (functional) endpoint is the registrational endpoint, is slow and highly variable, and requires large and long studies to demonstrate efficacy
- **Efficiency and cost of trials:** Structural endpoints could detect an earlier treatment effect which would allow for shorter (maybe 6-12 months) and smaller trials, accelerating clinical development
- **Increase confidence in treatment effect:** To have supplementary endpoints which increase confidence level of treatment benefits to patients

Goal: to link novel functional and structural endpoints to each other as well as tether to established primary objective of showing functional benefits

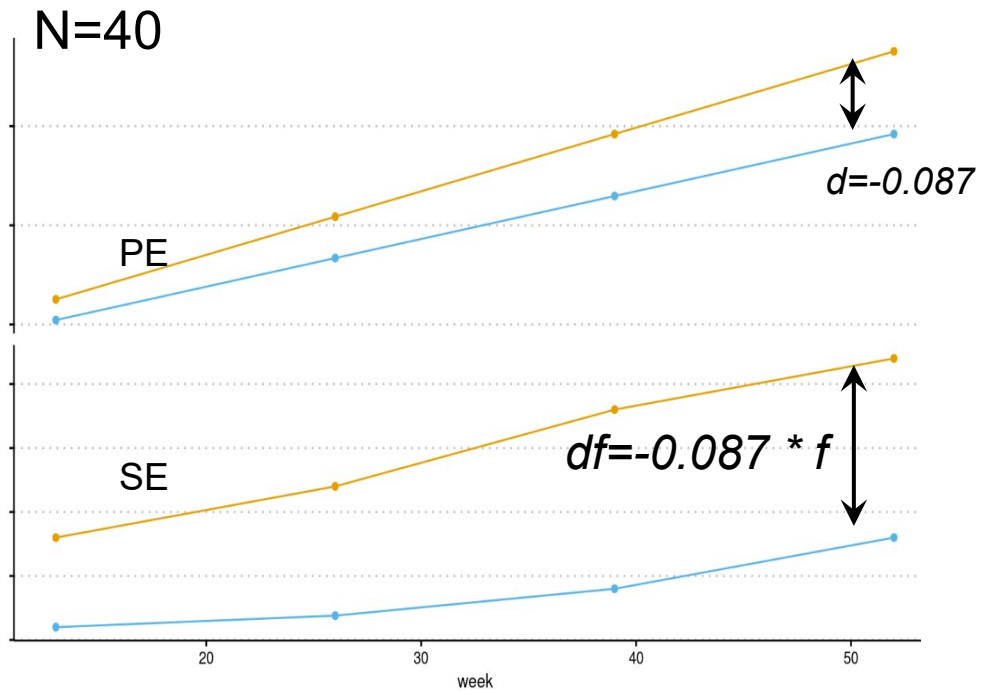
Simulation setup

Simulation parameter	Value
Sample size, n	80
Randomization ratio	treatment vs placebo = 1 vs 1
Enrolment	50 weeks to enroll all participants
Interim timing	All participants finished week 26
Schedule of activities	Week 13, 26, 39 and 52
PE*, Change from baseline (cfb) at week 52	• mean cfb in placebo = 0.32 • delta, d, difference in the mean cfb in treatment vs placebo = 0 (NULL), -0.08 (Moderate), -0.087 (TPP), and -0.096 (BEST) • Correlation over time, Autoregressive correlation, AR(1)=0.50
SE*, Change from baseline (cfb) at week 52	• mean cfb in placebo = 2.2 • delta, $d \times f$, difference in the mean cfb in treatment vs placebo = 0 (NULL), -0.08(Moderate) $\times$ 10, -0.087(TPP) $\times$ 10, and -0.096(BEST) $\times$ 10 • the relative difference $f = 10, 15$, or 20 • Correlation over time, Autoregressive correlation, AR(1)=0.50
Correlation (PE, SE)	0.80

Note: The lower the values in either endpoint, the better efficacy the treatment would provide in the endpoint. Homogeneous standard deviations in both treatment and placebo arms are assumed to be 0.18 and 1.2 for primary and secondary endpoints, respectively.

*PE - primary endpoint; SE - secondary endpoint.

➡ ~32 pts (40% of 80 total pts) finished W52



Statistical models (alpha = 10%, 1-sided)

A. Landmark Analysis with Available Data at Week 52

B. Parametric Longitudinal Modeling – Primary endpoint alone

C. Joint Longitudinal Modeling

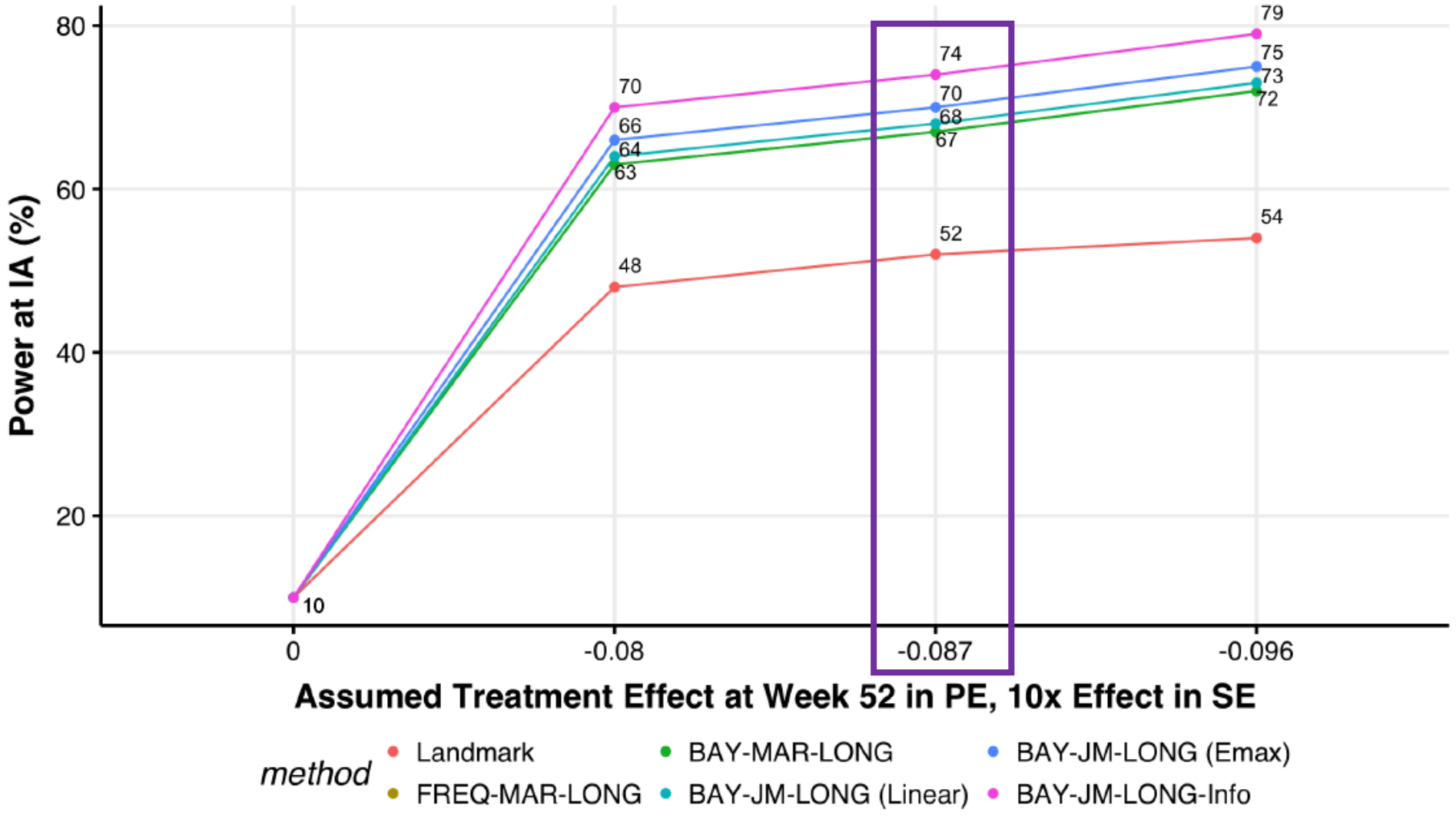
(1) Strongly informative

- Delta in PE CfB = f* x Delta in SE– PE CfB has to show improvement/worsening if SE CfB is showing improvement/worsening, so not only data in both endpoints impact on power

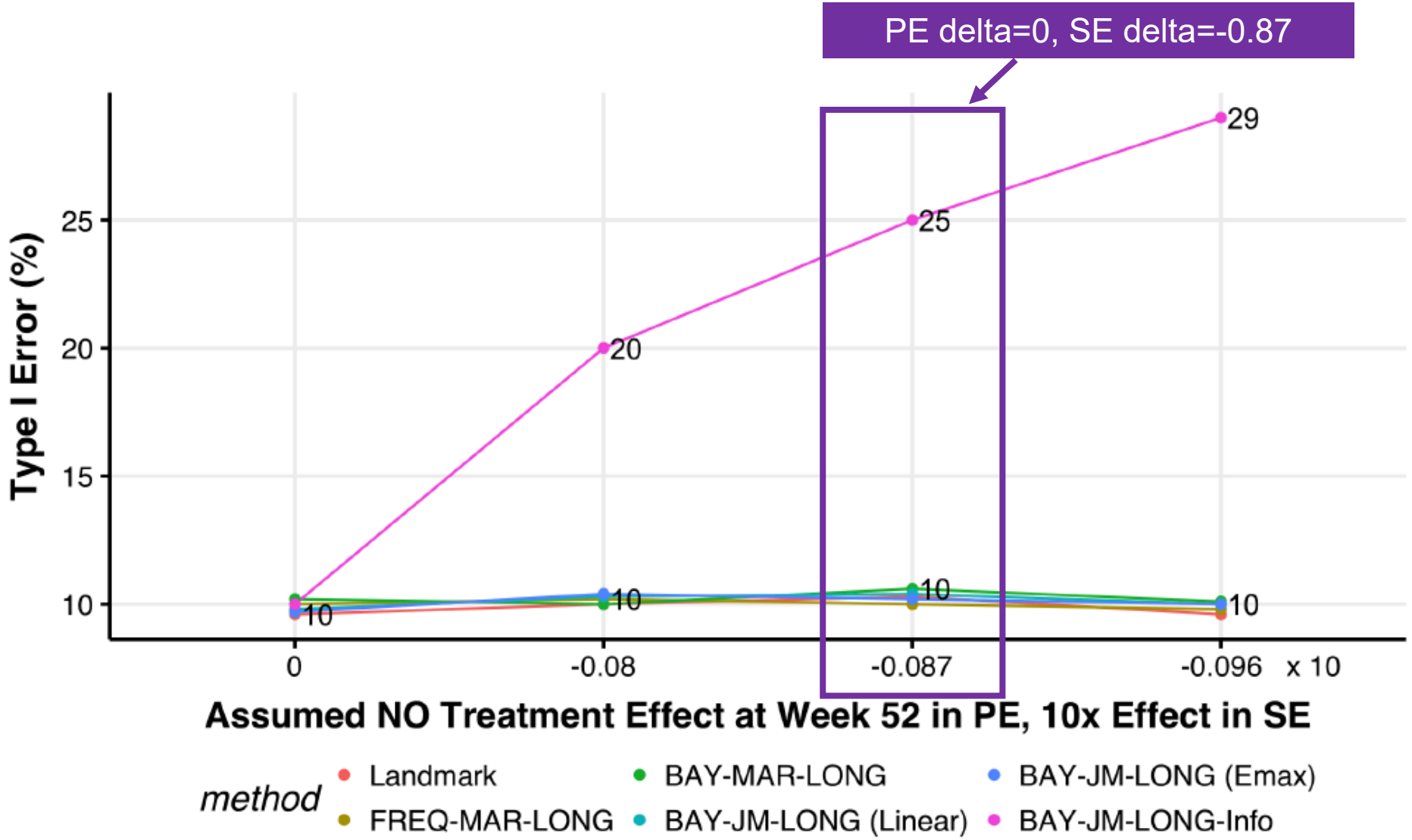
(2) Weakly informative

- Delta in PE CfB *"correlates with"* Delta in SE CfB – trajectories (intercepts and slopes) in CfB profiles are modelled jointly but highly non-informative priors are imposed, so data would dominate the statistical inference on power

Longitudinal modeling of PE CHG W52 gains ~20% power over landmark analysis
Longitudinal joint modeling (JM) yields ~6% power gain at inflated type I error



Joint modeling with strong prior inflates type I error from 10% to 25% if there is no effect in PE but full effect in SE



Joint modeling with strong prior inflates type I error from 10% to 25% if there is no effect in PE but full effect in SE

Recommendation: the use of "Longitudinal PE" method

- Relatively decent power of ~70% at IA ~20% gain from modeling only M12 data (power of ~50%)
- De-risk of inflating false positives by 2.5-fold compared with JM (Strong)

Method	Assumption	False positive	Comment
Landmark	✓ Primary endpoint of PE, no effect; ✓ SE has full effect	~10%	M12 data Only; PE Only
Longitudinal PE			All data; PE Only
JM (Weak)			All data; PE + SE; mean effects correlated via variability but directly through magnitude itself
JM (Strong)		25%	All data; PE + SE; mean effects are functionally connected and modelled

Application to Dose Response (Dummy/Simulated data)

Leveraging rich data across endpoints for understanding treatment effects across doses

Model setup:

- Endpoints
  - Y_{ijk} - endpoint k in patient i at visit j
  - $Y_{ijk} \sim N(e_0[z_i,k] + emax[z_i,k] \cdot \frac{visit_j}{visit_j + et50[z_i,k]}, \mathrm{V}^{\mathrm{k}}_{J\times J})$
  - $\mathrm{V}^{k}_{J\times J}$ has *AR(1)* auto-correlation structure
- Priors
  - $e_0[z_i, 1{:}K] \sim N(rep(0,K), \Omega^1_{K\times K})$
  - $emax[z_i, 1{:}K] \sim N(rep(0,K), \Omega^2_{K\times K})$
  - $et50[z_i, 1{:}K] \sim N(rep(5,K), \Omega^3_{K\times K})$
  - $\Omega^1_{K\times K} \sim Wishart(diag(K), K)$
  - $\Omega^2_{K\times K} \sim Wishart(diag(K), K)$
  - $\Omega^3_{K\times K} \sim Wishart(diag(K), K)$

Data inputs:

Y_{ij1} - ID Z wk0 wk2 wk4 wk8 wk16 wk24

Y_{ij2} - ID Z wk0 wk2 wk8 wk16 wk24

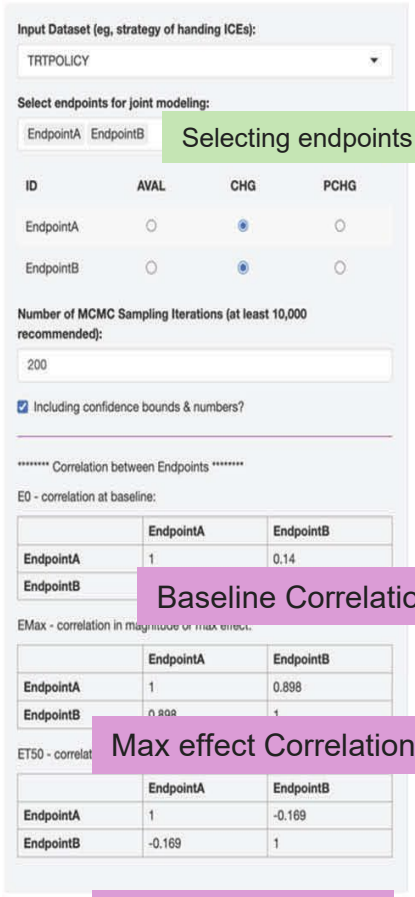
...

Y_{ijK} - ID Z wk8 wk16 wk24

Outputs:

- Auto-correlation for endpoint k
- Correlation between endpoints through Emax model parameters
- Added value in terms of bias and variance – to be explored

Output preview:

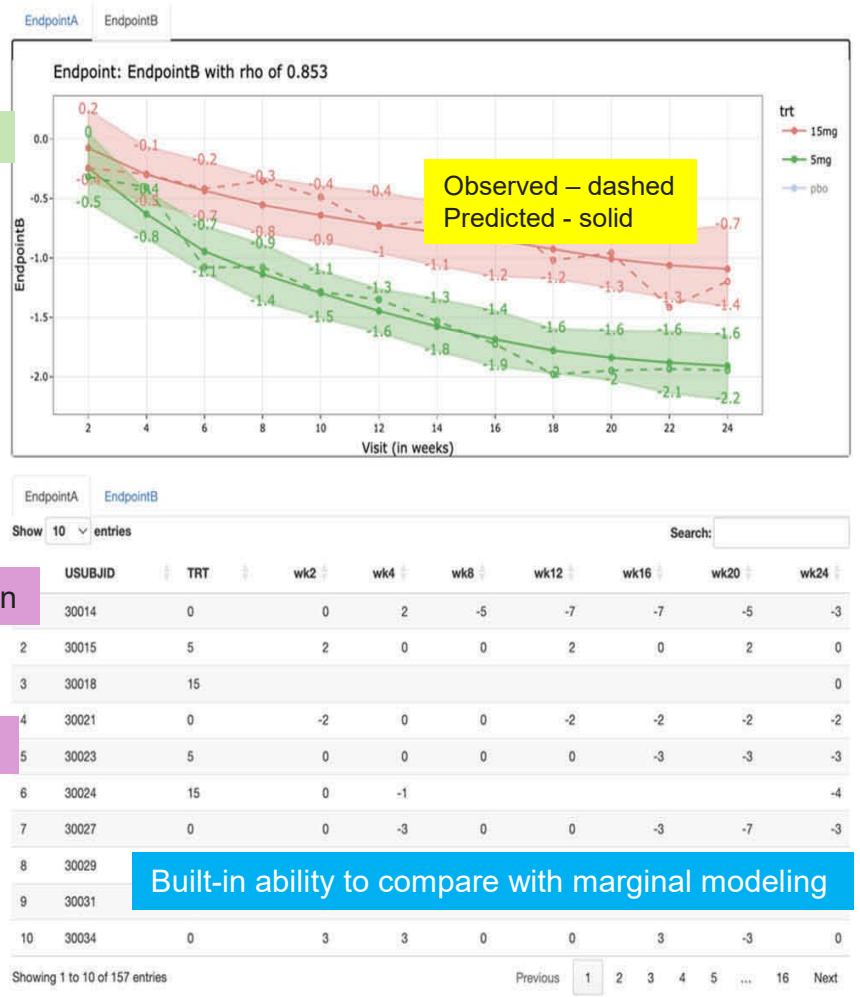


Selecting endpoints

Baseline Correlation

Max effect Correlation

Progression speed



Observed – dashed
Predicted - solid

Built-in ability to compare with marginal modeling

Input Dataset (eg, strategy of handing ICEs):

TRTPOLICY

Select endpoints for joint modeling:

EndpointA EndpointB EndpointC EndpointD

ID	AVAL	CHG	PCHG
EndpointA	○	◉	○
EndpointB	○	◉	○
EndpointC	◉	○	○
EndpointD	○	◉	○

Selecting endpoints

Number of MCMC Sampling Iterations (at least 10,000 recommended):

10000

☑ Including confidence bounds & numbers?

******** Correlation between Endpoints ********

E0 - correlation at baseline:

	EndpointA	EndpointB	EndpointC	EndpointD
EndpointA	1	0.139		
EndpointB	0.139	1	-0.022	-0.044
EndpointC	-0.01	-0.022	1	-0.004
EndpointD	0.06	-0.044	-0.004	1

Baseline Correlation

EMax - correlation in magnitude or max effect:

	EndpointA	EndpointB	EndpointC	EndpointD
EndpointA	1	0.932		
EndpointB	0.932	1		
EndpointC	-0.65	-0.618	1	0.382
EndpointD	-0.6	-0.599	0.382	1

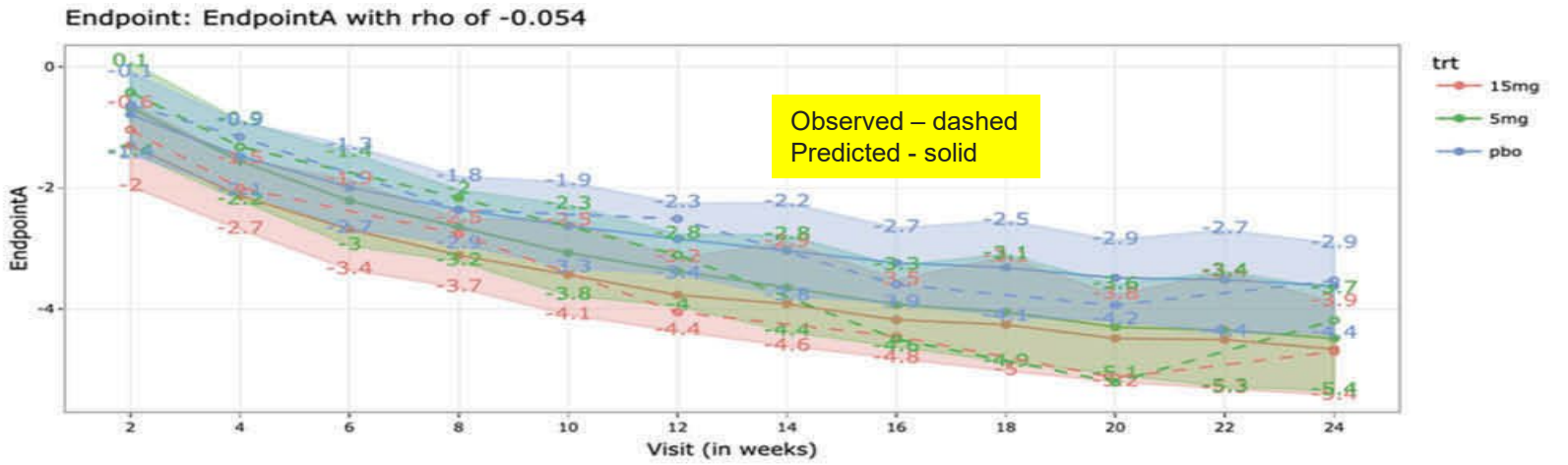
Max effect Correlation

ET50 - correlation in progression of effect:

	EndpointA	EndpointB	EndpointC	EndpointD
EndpointA	1	0.561	0.568	-0.451
EndpointB	0.561	1		
EndpointC	0.568	0.562		
EndpointD	-0.451	-0.459	-0.486	1

Progression speed

EndpointA EndpointB EndpointC EndpointD



EndpointA EndpointB EndpointC EndpointD

Show 10 entries Search:

	USUBJID	TRT	wk2	wk4	wk8	wk12	wk16	wk20	wk24
1	30014	0	0	2	-5	-7	-7	-5	-3
2	30015	5	2	0	0	2	0	2	0
3	30018	15							0
4	30021	0	-2	0	0	-2	-2	-2	-2
5	30023	5	0	0	0	0	-3	-3	-3
6	30024	15	0	-1					-4
7	30027	0	0	-3	0	0	-3	-7	-3
8	30029	5	-3	-3	-3				
9	30031	15	0	0	5		0	-4	-4
10	30034	0	3	3	0	0	3	-3	0

Showing 1 to 10 of 157 entries Previous 1 2 3 4 5 … 16 Next

Built-in ability to compare with marginal modeling

Results from simulated data

Comparison of multivariate joint longitudinal modeling and marginal (independent) longitudinal modeling (marginal vs joint) across dose groups.

Endpoint	Type	Estimates at Week 24, mean (95% CI)					
		Placebo		Low Dose		High Dose	
		JM	Marginal	JM	Marginal	JM	Marginal
EndpointA	Continuous	-3.5 (-4.3,-2.8)	-4.1 (-4.8, -3.3)	-4.4 (-5.2,-3.6)	-5.0 (-6.1, -4.1)	-4.6 (-5.4,-3.8)	-5.3 (-6.2, -4.4)
EndpointB	Continuous	-1.4 (-1.6,-1.2)	-1.5 (-1.8,-1.2)	-1.9 (-2.1,-1.6)	-2.0 (-2.3,-1.7)	-1.0 (-1.3,-0.8)	-1.1 (-1.4,-0.8)
EndpointC	Binary	0.5 (0.4,0.6)	0.5 (0.4,0.6)	0.4 (0.3,0.5)	0.4 (0.3,0.5)	0.4 (0.3,0.55)	0.4 (0.3,0.5)
EndpointD	Continuous	-0.2 (-0.9,1.4)	-0.2 (-1.5,0.9)	1.6 (0.4,3.0)	1.6 (0.2,3.1)	0.3 (-0.9,1.5)	0.0 (-1.2,1.0)

Estimated endpoint effect for high dose shows narrower CIs compared with marginal modeling

Results from simulated data

Correlations between endpoints at baseline, magnitude or max effect, and progression of the treatment effects

******** Correlation between Endpoints ********

E0 - correlation at baseline:

	EndpointA	EndpointB	EndpointC	EndpointD
EndpointA	1	0.139	-0.01	0.06
EndpointB	0.139	1	-0.022	-0.044
EndpointC	-0.01	-0.022	1	-0.004
EndpointD	0.06	-0.044	-0.004	1

EMax - correlation in magnitude or max effect:

	EndpointA	EndpointB	EndpointC	EndpointD
EndpointA	1	0.932	-0.65	-0.6
EndpointB	0.932	1	-0.618	-0.599
EndpointC	-0.65	-0.618	1	0.382
EndpointD	-0.6	-0.599	0.382	1

ET50 - correlation in progression of effect:

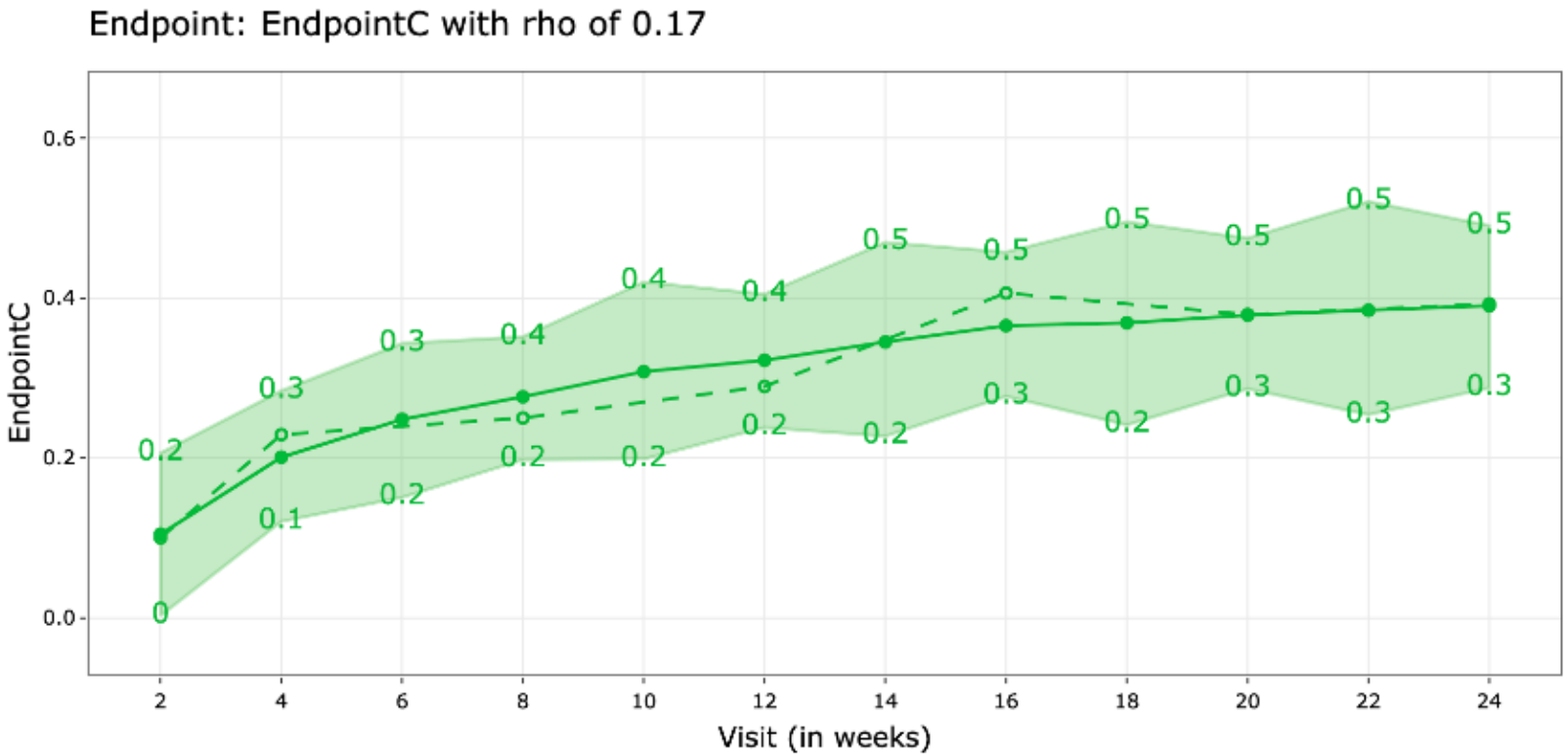
	EndpointA	EndpointB	EndpointC	EndpointD
EndpointA	1	0.561	0.568	-0.451
EndpointB	0.561	1	0.562	-0.459
EndpointC	0.568	0.562	1	-0.486
EndpointD	-0.451	-0.459	-0.486	1

Summary:

- ❑ Weak correlations at baseline between A/B/C and MoA with the strengths of 0.139, -0.01, and 0.06, respectively
- ❑ Moderate to strong correlations at the magnitude and progression rates
  - A and B are strongly related at magnitude, and moderately related at progression speed
  - A and D are moderately & negatively correlated at magnitude & progression speed over time

Example result from one endpoint

Marginal estimates of treatment effects over time in the *Endpoint C*. Comparison of observed (dashed lines) and estimated (95%CI, solid line and shadow area) treatment effects for the low dose group



<u>Within endpoint correlation</u>: The <u>SIMULATED</u> data shows there is a weakly correlation (0.17) over time in Endpoint C

Remarks

- Joint modeling provides gains of efficiency in terms of treatment effect estimation
- Properly setting up the joint model could uncover relationship among endpoints if there are different sorts of correlations within / between endpoints (e.g., autocorrelations within each endpoint, correlations at baseline, correlations in progression speed, correlation in magnitudes of endpoints)
- Leveraging all sources of data through joint modeling would help both trial monitoring (dose response, endpoint establishment) and design planning (potential power gain & estimates).
- Inhouse App developed for better communication purposes

Future directions

Considering clinical implications, or meaningful interpretations, how to address challenges *like*

- ❑ Early stop for efficacy, futility
- ❑ Biomarker-guided safety review
- ❑ Benefit risk analysis using utility scores
- ❑ Extension to time-to-event endpoints

Additional Points to Consider

Model selection

Check model fits by visualization

Statistical testing & model comparisons

Diagnosis

Diagnosis of model convergence

Diagnosis of model assumptions

Thank you!!

Any Questions?

Johnson&Johnson
Innovative Medicine