Methods: Mind the Gap
Webinar Series

Power Analyses to Plan Idiographic Clinical Trials, Illustrated for Prevention and Rare Diseases

Presented by:

Ty Ridenour, Ph.D., M.A.
Stephen Tueller, Ph.D., M.S.
RTI International

National Institutes of Health
Office of Disease Prevention

Power Analyses to Plan Idiographic Clinical Trials, Illustrated with Prevention and Rare Diseases

Ty A. Ridenour & Stephen J. Tueller

Funded by NCATS (R21 TR002402) and RTI IR&D Innovation Grants

PersonAlytics™ Team and Website: https://personalytics.rti.org/

Outline

- When to (not) use Idiographic Clinical Trials (ICTs)
- What are ICTs?
- Illustrations of ICTs
- Recent developments
- Power analysis software

Motivations for Idiographic Clinical Trials

Small population or sample

In-the-field research

When withholding treatment is unethical

Active ingredients / processes

Precision treatment

What works for whom

Rapid program evaluation

Underserved populations

When to Generally (not) Use ICTs

ICTs Generally Strong For:

- N = 1 results ("impact")
- Comparative effectiveness
- All participants get novel treatment
- Engagement / attrition
- Intrapersonal processes / mechanisms
- Real world effectiveness
- "Active ingredients" research
- Small population efficacy

ICTs Generally Limited For:

- Large population efficacy
- Acute illnesses
- Change in traits / personality
- Few "waves"
- Surveys / prevalence
- Long interviews / questionnaires
- Note: Stigma among methodologists

What are Idiographic Clinical Trials?

Within-subject Experimental Designs + **Time Series Data** + **Stochastic Analysis for Small N** = **Idiographic Clinical Trials**

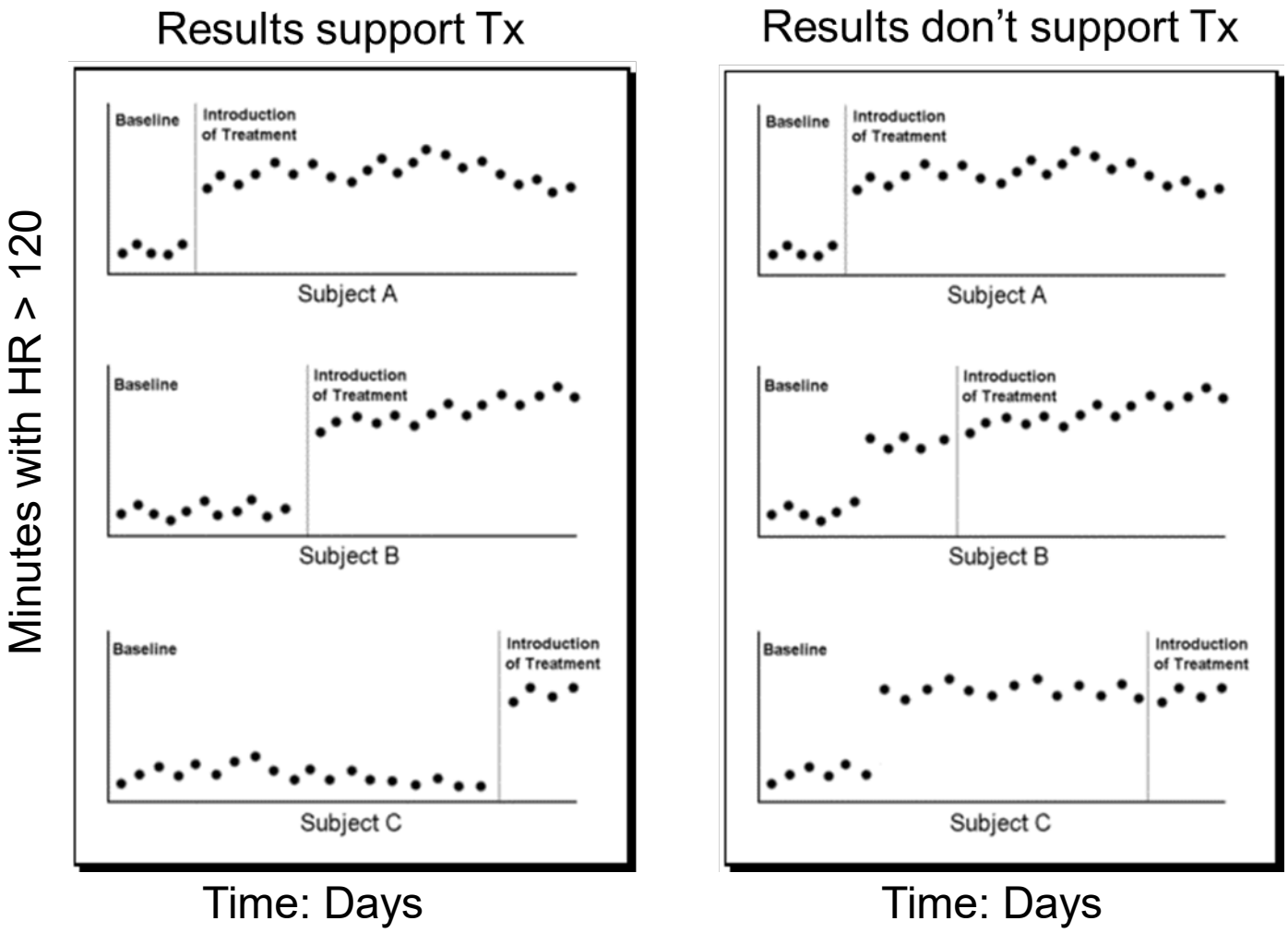
- Each participant get 2+ phases
- Logical, flexible, causal designs

- Per phase
- Sensitive to change
- Valid for detecting change

- Can yield aggregates
- Efficacy-like output
- Resolves sources of bias in WSEDs

Within-subject Experimental Designs

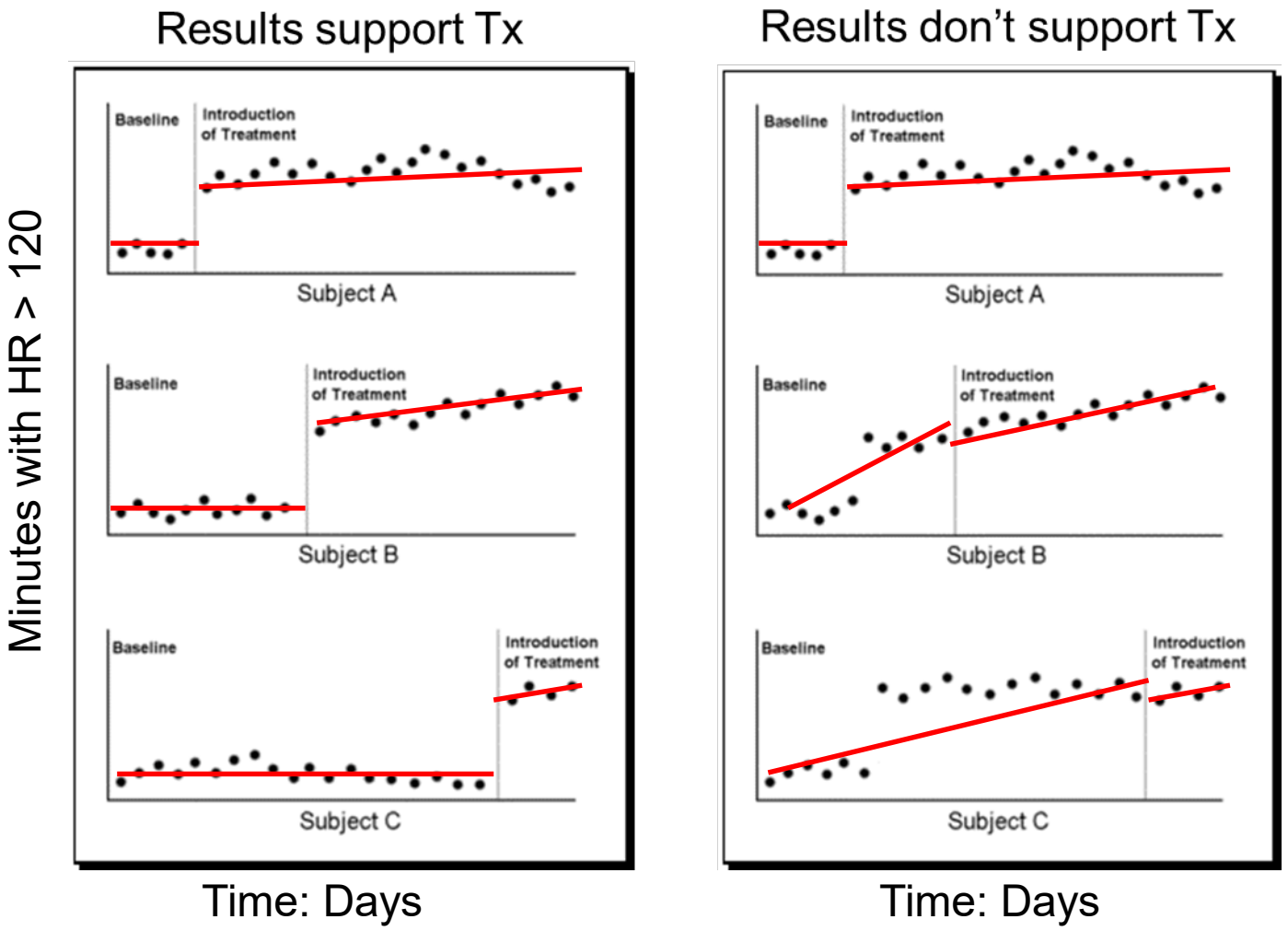
Most Common: Multiple Baseline Design



From: AllPsych; https://allpsych.com/research-methods/singlesubjectdesign/multiplebaselines/#.Vd30PvlVhBe

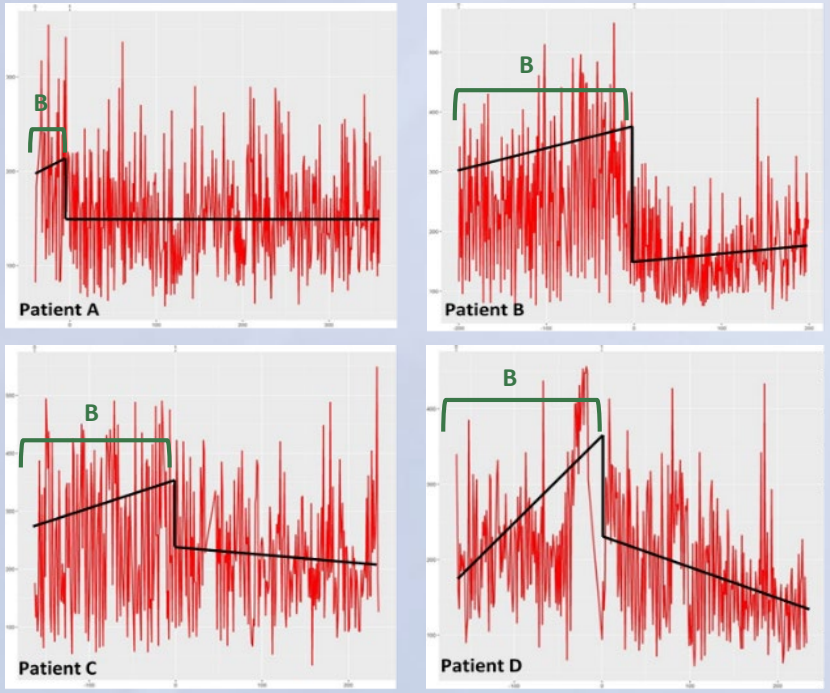
Within-subject Experimental Designs

Most Common: Multiple Baseline Design

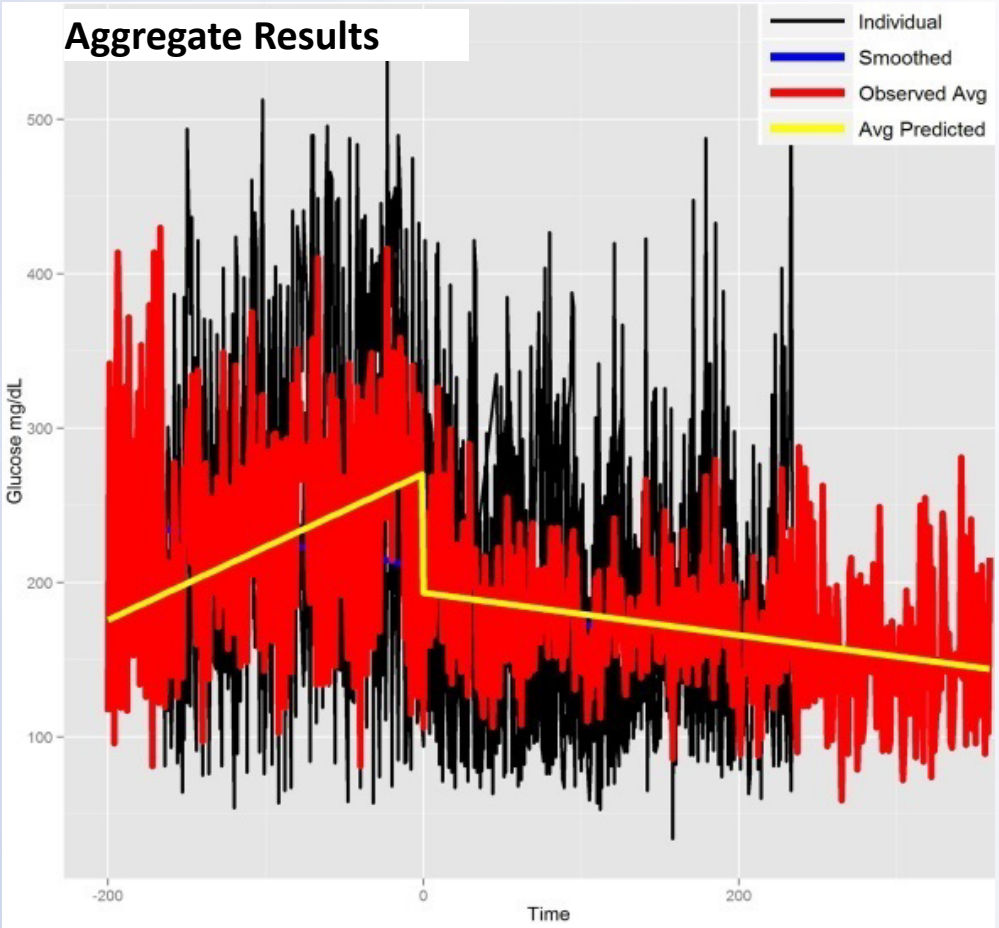


From: AllPsych; https://allpsych.com/research-methods/singlesubjectdesign/multiplebaselines/#.Vd30PvlVhBe

Illustration 1: Rigorous Pilot Study



Note: Y-axis is blood glucose in mg/dL. **B** = baseline phases. The treatment instant impact (without slope) in mg/dL is -49.0 for A, -152.9 for B, -45.0 for C, and -73.0 for D.



Note: Average decrease in glucose immediately following Manual Pancreas = 77.13 mg/dL ($p < .001$). Smoothed model not shown.

From: Ridenour et al., 2013

Analytic Strategy: Intensive Hierarchical Regression

$$y_{ij} = \pi_{0i} + \pi_{1i} Time_j + \pi_2 \left(Time_j \times Phase_j \right) + \pi_3 Phase_j + \varepsilon_{ij}$$

Where:

y_{ij} represents outcomes for individual i at time j

π_{0i} represents random intercepts

$\pi_{1i}\ Time_j$ represents random slopes

$Phase_j$ is dummy coded to estimate the effect of time separately by phase

$\pi_2 \left(Time_j \times Phase_j \right)$ is a fixed effect of time

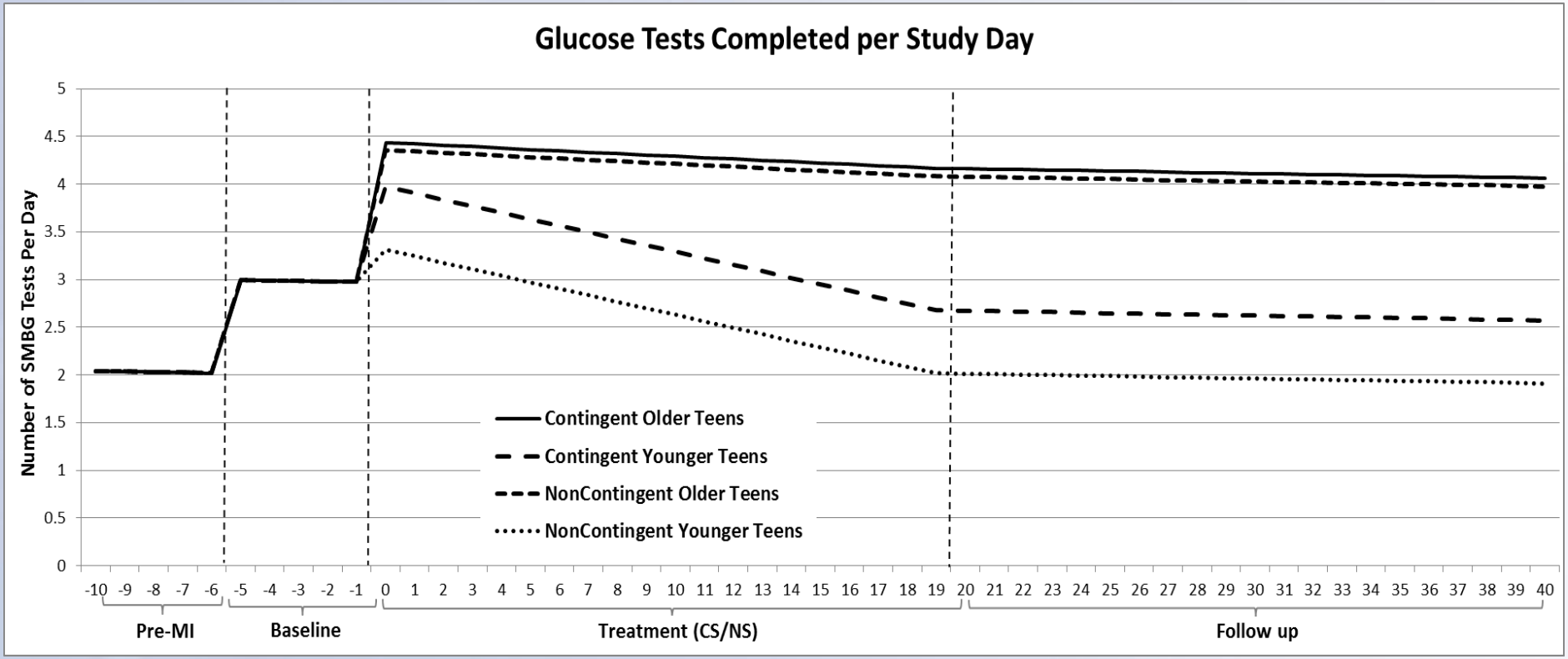
$\pi_3 Phase_j$ a fixed effect of difference in intercepts among phases

ε_{ij} is residual variance term

Model assumes that during baseline the mean intercept = 0 and mean slope = 0; autoregression in data has been parsed out using the appropriate error covariance structure; and error is uncorrelated with random effects.

Can add term(s) to test subgroup differences and analyze covariates.

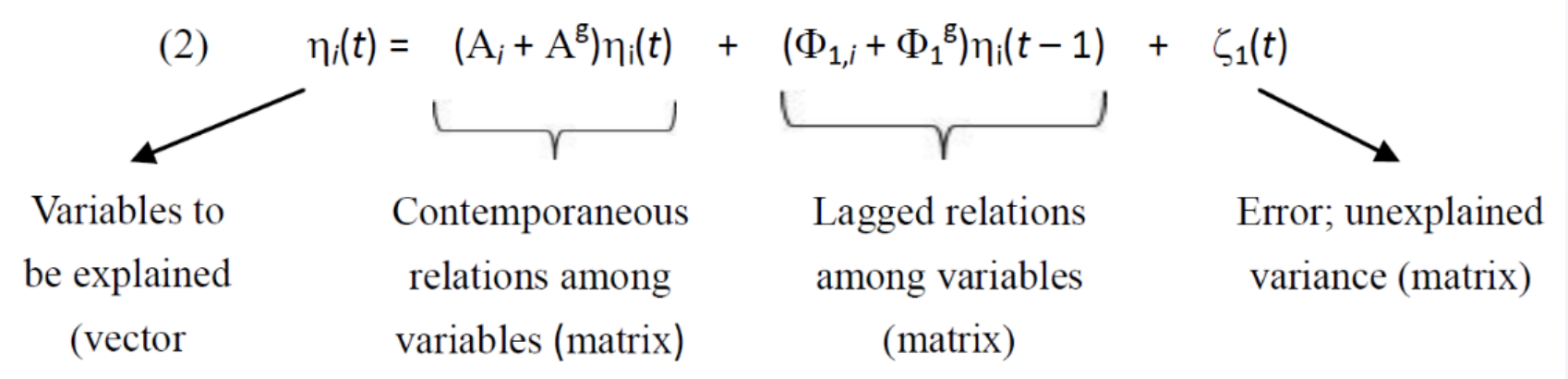
Illustration 2: Comparative Effectiveness Research



Daily Tests = 1.9885 - 0.00501 (per day) + 0.9805 (effect of MI) + 1.3240 (change in intercept at Treatment phase) - 0.06317 (per day of Treatment phase) + 1.0430 (additional intercept change for older teens during Treatment phase) + 0.6598 (additional intercept for CS) – 0.05378 (per day of Treatment phase for younger teens)

From: Raiff et al., 2016

Analytic Strategy: Unified SEM



Where:

$\eta_i(t)$ are the variables to be "explained" for individual *i*

$(A_i + A^g)\eta_i(t)$ is a matrix of contemporaneous covariations among variables

$(\Phi_{1,i} + \Phi_1{}^g)\ \eta_i(t\text{-}1)$ is a matrix of lagged covariations among variables

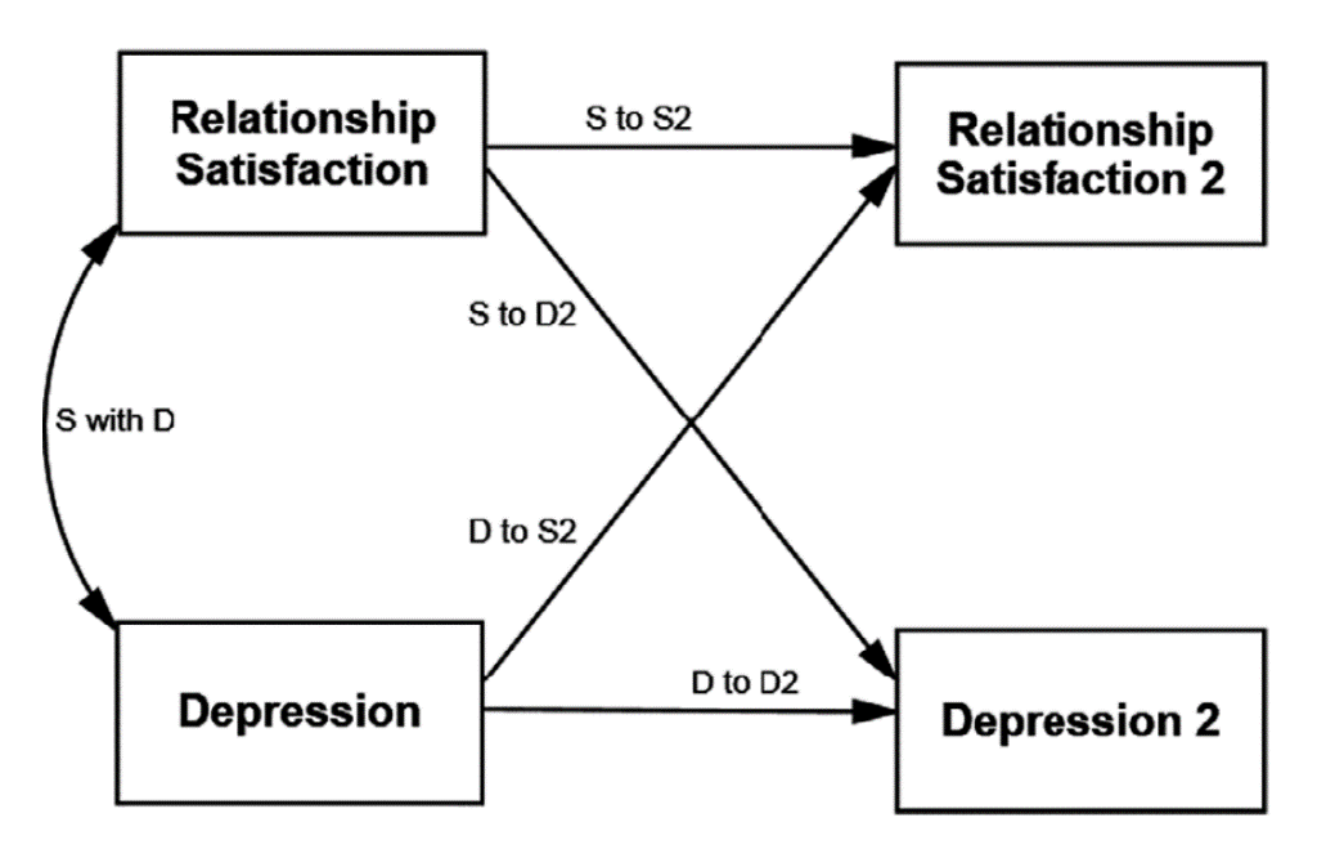
$\zeta_1(t)$ is an error matrix

Notation, assumptions, and modelling strategy are based on the Group Iterative Multiple Model Estimation (GIMME) programs.

From: Beltz et al., 2016; Gates et al., 2012

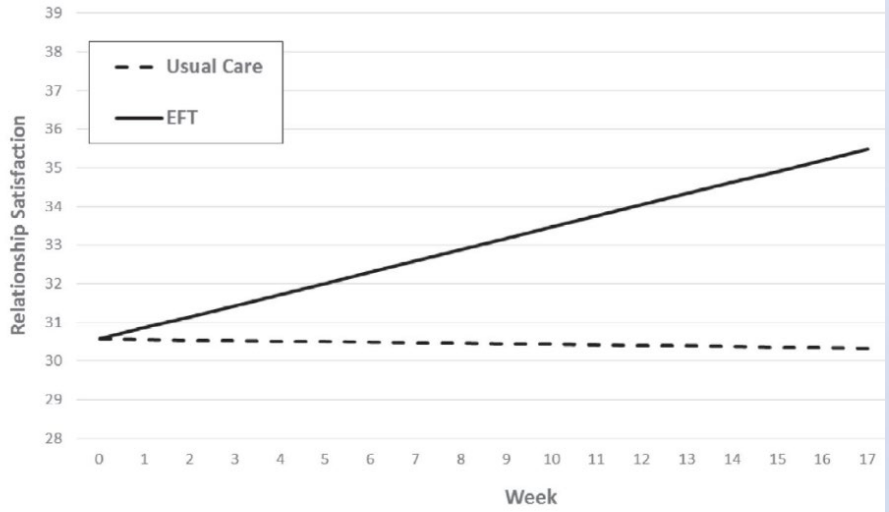
Illustration 3: Testing Mechanisms of Action

Hypothesized model of Emotion Focused Therapy outcomes



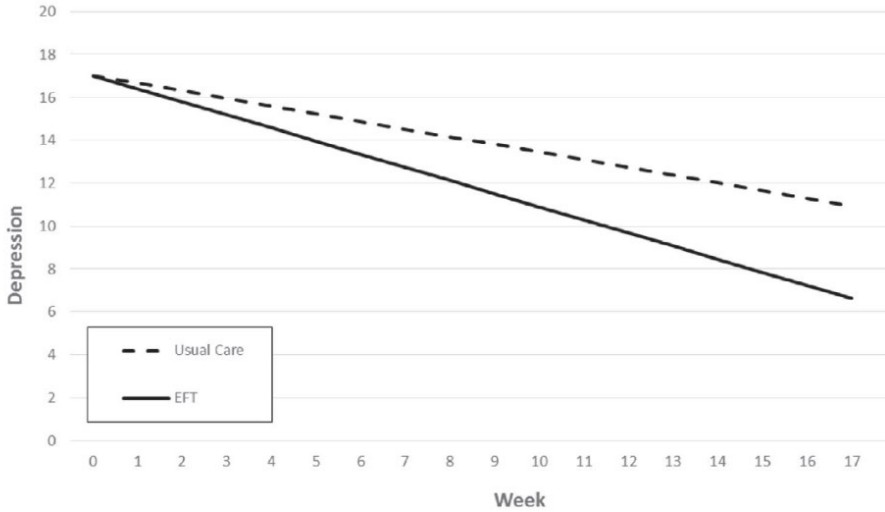
From: Wittenborn et al., 2019; Ridenour et al., 2016

Illustration 3: Outcomes for the Men



Satisfaction = 30.57 + 0.29 (Week) – 0.30 (Week*Treatment)

Figure 3. Mixed model trajectory analysis of relationship satisfaction.



Depression = 17.01 - 0.61 (Week) + 0.25 (Week*Treatment)

Figure 4. Mixed model trajectory analysis of depression.

From: Wittenborn et al., 2019

USEM: Testing of Fit to the Data

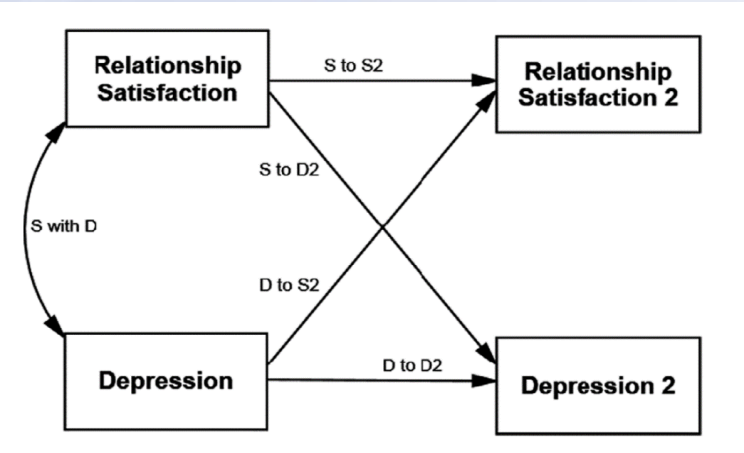


Table 3
Fit Statistics of Three Competing Subgroupings of Men

Path parameters fixed equal. . .	χ^2, *df*	AIC	BCC	LR χ^2, *df* vs. model 1
1. . . . across all participants	1199.09, 171	1277.1	1305.0	–, –
2. . . . within treatment arms	1183.18, 166	1271.2	1302.6	15.9, 5
3. . . .within each of 4 clusters	1124.27, 151*	1242.3*	1284.4*	58.9, 20*

Note. *df* = degrees of freedom; RMSEA = root mean square error of approximation; AIC = Akakie's Information Criterion; BCC = Brown-Cudeck Criterion; LR = likelihood ratio. Models 2 and 3 are not nested and thus were not compared using LR χ^2.
*The best fitting model indicated by the fit statistic.

Source:
https://www.researchgate.net/publication/268146083_Demonstration_of_Two_Traditional_Statistical_Techniques_for_Use_with_Small_Sample_within-Person_Experiments_Unified_Structural_Equations_Modeling_and_Mixed_Model_Trajectory_Analysis

USEM: Differential Treatment Responses

Table 4
Standardized Path Coefficients of the Four-cluster Solution for Men

Aggregate estimates	ID	Autocorrelation		Cross-lag paths		Cluster path characteristics	Study arm
		$S \rightarrow S_2$	$D \rightarrow D_2$	$S \rightarrow D_2$	$D \rightarrow S_2$		
Cluster 1:	20	−0.02	0.64	−0.30	−0.96	Autocorrelation in depression only; Granger causality from depression to satisfaction	UC
$S \rightarrow S_2 = 0.03$;	25	−0.05	0.71	−0.13	−0.56		UC
$D \rightarrow D_2 = 0.26$;	27	−0.08	0.71	0.04	−0.54		EFT
$S \rightarrow D_2 = -0.37$;							
$D \rightarrow S_2 = -0.35$							
Cluster 2:	11	1.09	0.01	−0.69	0.26	Autocorrelation in satisfaction only; Granger causality from satisfaction to depression; lesser sequence from depression to satisfaction	EFT
$S \rightarrow S_2 = 0.77$;	21	0.50	−0.09	−0.83	−0.40		EFT
$D \rightarrow D_2 = 0.02$;	26	0.47	0.05	−0.39	−0.28		UC
$S \rightarrow D_2 = -0.71$;							
$D \rightarrow S_2 = -0.05$							
Cluster 3:	8	0.16	0.50	0.31	−0.24	Moderate autocorrelation for depression; small-to-nil cross-lagged correlations	UC
$S \rightarrow S_2 = 0.43$;	15	0.16	0.50	0.00	−0.20		EFT
$D \rightarrow D_2 = 0.31$;	22	−0.33	0.33	0.11	−0.23		EFT
$S \rightarrow D_2 = -0.17$;							
$D \rightarrow S_2 = -0.14$							
Cluster 4:	2	0.65	0.64	−0.32	−0.20	Large autocorrelations for depression and satisfaction; moderate-to-nil cross-lagged correlations	EFT
$S \rightarrow S_2 = 0.40$;	3	0.86	0.76	−0.19	0.01		EFT
$D \rightarrow D_2 = 0.50$;	16	0.54	0.90	−0.04	−0.45		UC
$S \rightarrow D_2 = -0.17$;	23	0.63	0.91	−0.02	0.19		UC
$D \rightarrow S_2 = -0.14$	28	0.55	0.76	0.12	0.05		EFT

Note. S = relationship satisfaction; D = depression. Model parameters of one participant (ID 24) did not fit into any of the clusters; they were −0.71, 0.07, 2.23, and −0.09, respectively.

Source:
https://www.researchgate.net/publication/268146083_Demonstration_of_Two_Traditional_Statistical_Techniques_for_Use_with_Small_Sample_within-Person_Experiments_Unified_Structural_Equations_Modeling_and_Mixed_Model_Trajectory_Analysis

From: Wittenborn et al., 2019

(Some) Recent Advances for ICTs

Understanding ICT outcomes as “factuals” & “counterfactuals”: Daza et al., 2018

Simulations to inform study design: Blackston et al., 2019; Duan et al., 2013; Percha et al., 2019; Tueller et al., 2022

Understanding patient preferences for study designs (by illness): Cheung et al., 2020; Sacristán et al., 2021

Alternative designs and analytic strategies: Howe et al., 2010; Liao et al., 2021; Nahum-Shani et al., 2015

Causal mediation analysis: Special issue (2022) in *Evaluation & the Health Professions*: Miočević et al., 2022

Some Resources

Stats-of-1: Inference for the Individual

References, links to useful tools

https://statsof1.org/resources/#sample-size--statistical-power

International Collaborative Network

https://www.nof1sced.org/

Single Case Design Masked Visual Analysis

Data visualization and sharing apps

https://singlecasemva.app/

Ksana Health data visualization apps

https://ksanahealth.com/ears/

Evolving Resource: *PersonAlytics*™

Statistical and power analysis programs to support ICTs

Automate certain analytic processes

Support simulation research

Provide GUI interface for users that don't code in R

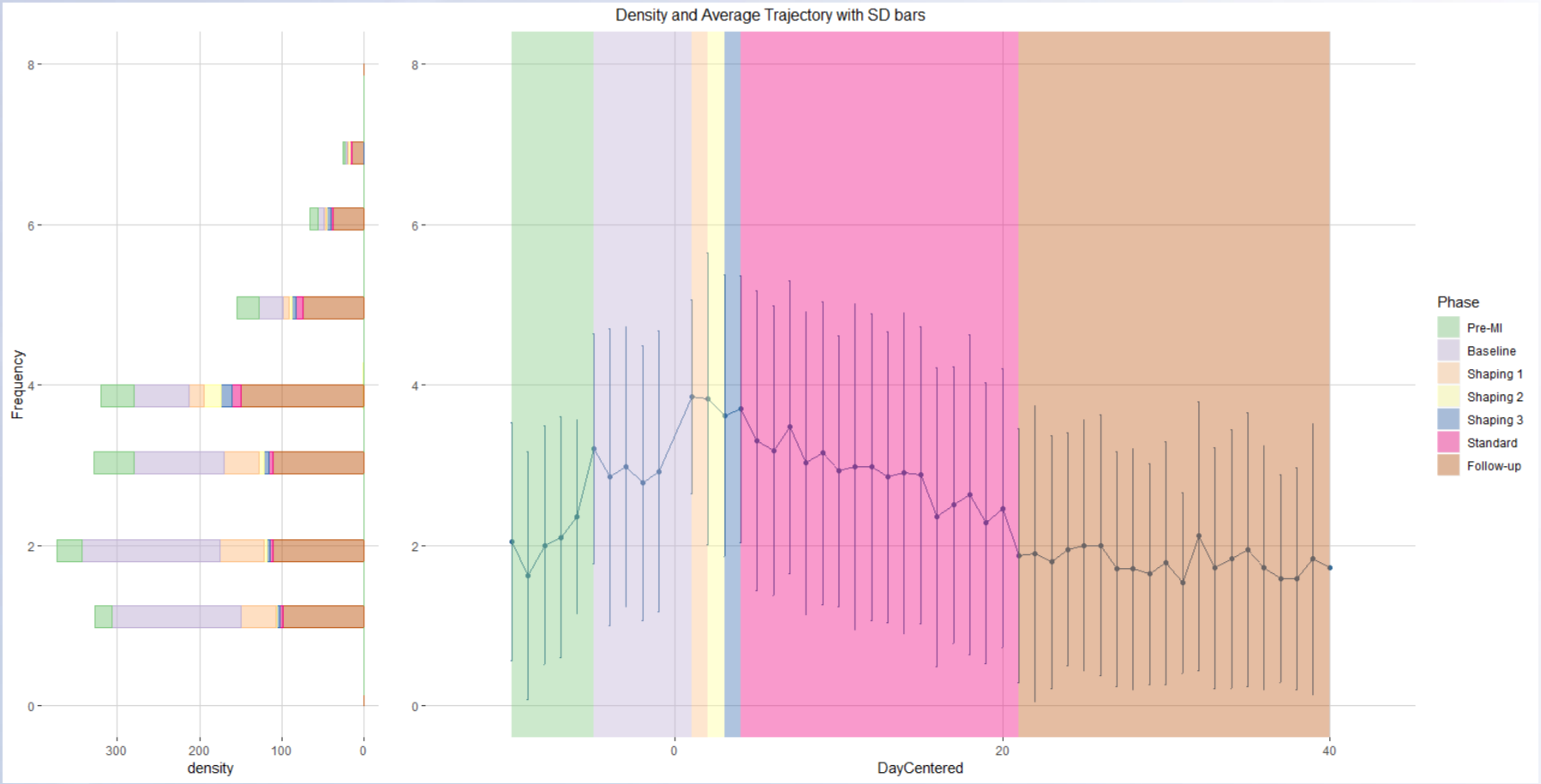
Evolve with methodological developments

Website: https://personalytics.rti.org/

PersonAlytics R Package

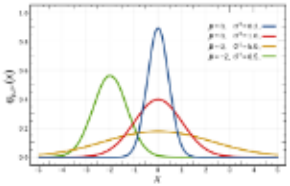
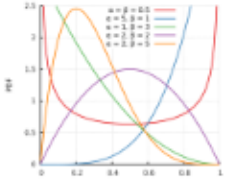
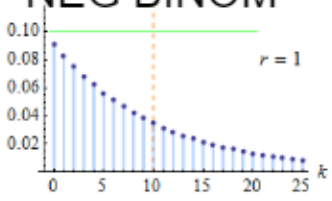
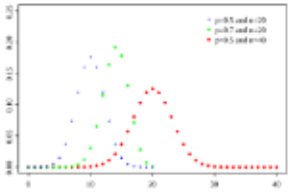
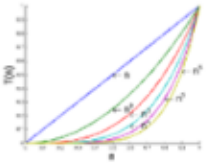
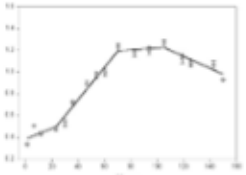
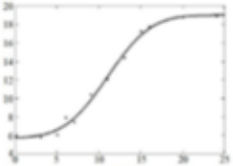
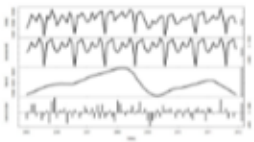
- Analytics for N-of-1 and small N intensive longitudinal designs, idiographic clinical trials (ICT), and interrupted time series https://github.com/ICTatRTI/PersonAlytics
- Single subject data: Linear ARMA models
- Small N data: Mixed effects models (MLM/HLM/GCM)
  - Linear mixed effects model
  - Generalized additive models for location, scale and shape (70+ distributions)
- Mixed effects modeling options
  - Standard MLM/HLM with polynomial orders of time ($time$, $time^2$, $time^3$)
  - Piecewise growth model
  - Simultaneous estimate of phase and group specific MLM/HLM/GCM
- Data visualization
- Finite population correction (FPC)

Visualizing ICT Data



Mixed Effects and Time Series Modeling for N=1, small N, and ICT

OUTCOME	=	FIXED EFFECTS	+	RANDOM EFFECTS	+	ERROR

	NORMAL	BETA	NEG BINOM	NEG BINOM
70+ OUTCOME DISTRIBUTIONS				
FIXED EFFECTS	EXPERIMENTAL CONDITION/ PHASE e.g. ABA study		DYNAMIC FACTORS e.g., weather, BP	IF N > 1 STATIC FACTORS e.g. Demographics
TRAJECTORY SHAPE		Polynomial 	Piecewise 	Logarithmic
ERROR COVARIANCE STRUCTURES	TIME SERIES e.g., ARMA, ARIMA		CORRELATION STRUCTURES e.g., compound symmetry, spherical, geospatial	

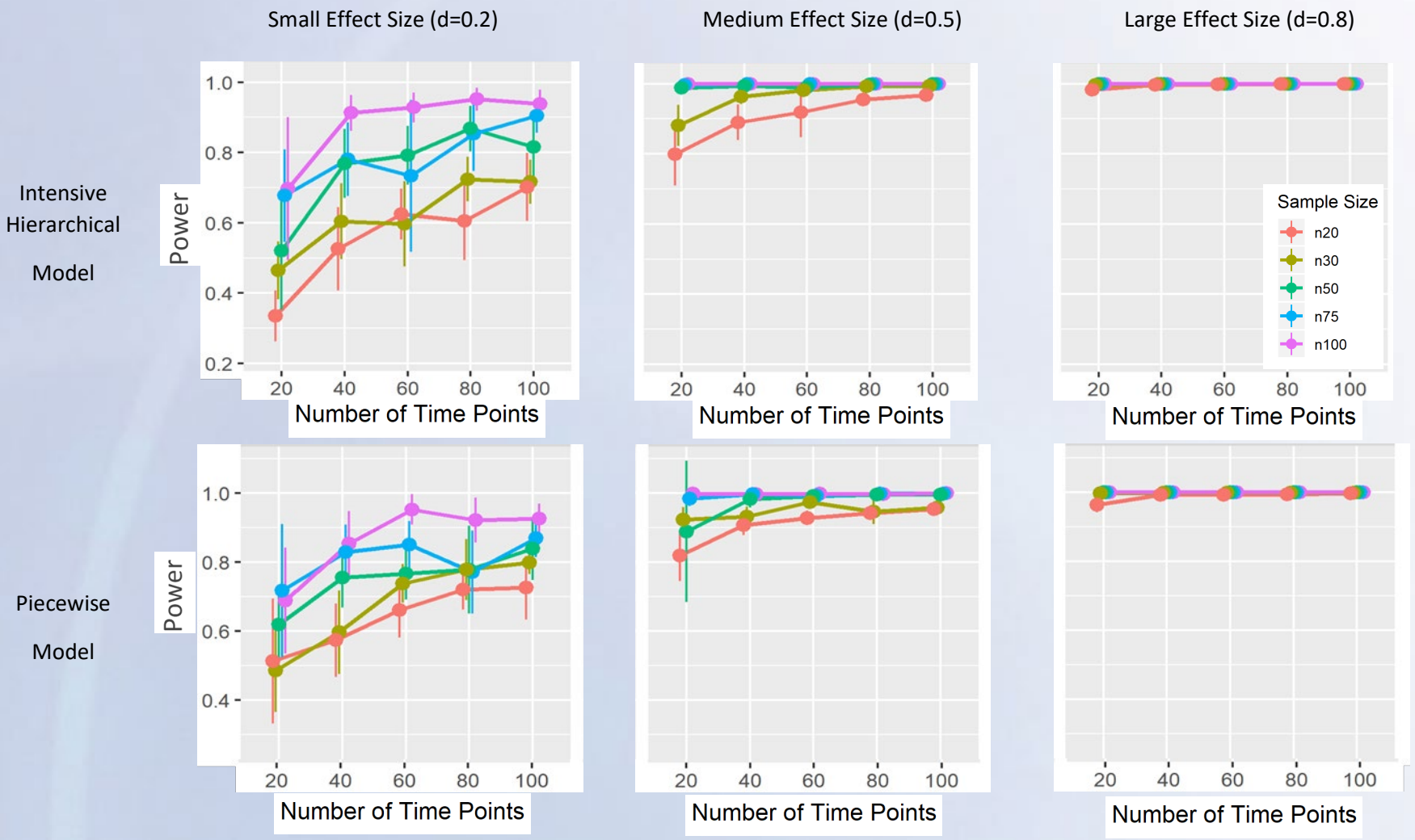
Modeling Process Automation Features

- Model selection using AIC or BIC
- All model selection uses ML, final model is fit with REML
- Automated tasks
  - Residual correlation structure selection
    - ARMA(p, q) for all possible combinations of p & q
    - User specified p & q
  - Time structure selection
    - Polynomial (time, $time^2$, $time^3$, etc.)
    - Pending feature: estimating polynomial time structure within each phase
  - Standardization of outcomes, predictors, or both
  - Centering of the time variable
  - Outcome distribution selection

PersonAlyticsPower R Package

- Power Analysis for N-of-1 and small N intensive longitudinal designs, idiographic clinical trials (ICT), and interrupted time series
- Simulation based power analysis for any number of phases or groups
- Binary and normal outcomes (other distributions in development)
- User inputs are average intercepts and slopes in each phase and each group with standardized effect size differences
- Web based GUI in development
- https://github.com/ICTatRTI/PersonAlyticsPower

PersonAlytics Power Analysis



Power Analyses to Plan Idiographic Clinical Trials, Illustrated with Prevention and Rare Diseases

Ty A. Ridenour, Ph.D., M.P.E.
Developmental Behavioral Epidemiologist, RTI International
TRidenour@rti.org

Stephen J. Tueller, Ph.D.
Senior Research Statistician, RTI International
STueller@rti.org

Funded by NCATS (R21 TR002402) and RTI IR&D Innovation Grants

References

Beltz, A. M., Wright, A. G., Sprague, B. N., & Molenaar, P. C. (2016). Bridging the nomothetic and idiographic approaches to the analysis of clinical data. *Assessment*, *23*, 447-458.

Blackston, J. W., Chapple, A. G., McGree, J. M., McDonald, S., & Nikles, J. (2019) Comparison of aggregated N-of-1 trials with parallel and crossover randomized controlled trials using simulation studies. In *Healthcare*, *7*, 137. MDPI.

Cheung, Y. K., Wood, D., Zhang, K., Ridenour, T. A., Derby, L., St. Onge, T., Duan, N., Duer-Hefele, J., Davidson, K.W., Kronish, I. and Moise, N., (2020). Personal preferences for personalised trials among patients with chronic diseases: an empirical Bayesian analysis of a conjoint survey. *BMJ Open, 10*, e036056.

Daza, E. J. (2018). Causal analysis of self-tracked time series data using a counterfactual framework for N-of-1 trials. *Methods of Information in Medicine, 57*(S 01), e10–e21.

Duan, N., Kravitz, R. L., & Schmid, C. H. (2013). Single-patient (N-of-1) trials: A pragmatic clinical decision methodology for patient centered comparative effectiveness research. *Journal of Clinical Epidemiology, 66*, S21–S28.

Gates, K. M., & Molenaar, P. C. (2012). Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *NeuroImage*, *63*, 310-319.

Howe, G. W., Beach, S. R., & Brody, G. H. (2010). Microtrial methods for translating gene-environment dynamics into preventive interventions. *Prevention Science, 11*, 343–354.

Liao, Z., Qian, M., Kronish, I. M., & Cheung, Y. K. (2021). Analysis of N-of-1 trials using Bayesian distributed lag model with autocorrelated errors. *arXiv preprint arXiv:2112.13991*.

Miočević, M., Moeyaert, M., Mayer, A., & Montoya, A. K. (2022). Causal Mediation Analysis in Single Case Experimental Designs: Introduction to the Special Issue. *Evaluation and the Health Professions*, *45*, 3-7.

References (cont.)

Nahum-Shani, I., Hekler, E. B., & Spruijt-Metz, D. (2015). Building health behavior models to guide the development of just-in-time adaptive interventions: A pragmatic framework. *Health Psychology*, *34*(S), 1209.

Percha, B., Baskerville, E. B., Johnson, M., Dudley, J. T., & Zimmerman, N. (2019). Designing Robust N-of-1 studies for precision medicine: Simulation study and design recommendations. *Journal of Medical Internet Research, 21*, e12641.

Raiff BR, Barry VB, Jitnarin N, Ridenour TA. Internet-based incentives increase blood glucose testing with a non-adherent, diverse sample of teens with Type 1 Diabetes Mellitus. *Translational Behavioral Medicine* 2016, 6, 179-188.

Ridenour TA, Pineo TZ, Maldonado-Molina MM, Hassmiller-Lich, K. (2013). Toward idiographic research in prevention science: Demonstration of three techniques for rigorous small sample research. *Prevention Science*, *14*, 267-278.

Ridenour TA, Wittenborn AK, Raiff BR, Benedict N., Kane-Gill S. (2016). Illustrating idiographic methods for translation research: Moderation effects, natural clinical experiments, and complex treatment-by-subgroup interactions. *Translational Behavioral Medicine*, *6*, 125-134.

Sacristán, J. A., & Knottnerus, J. A. (2021). Individual point-of-care trials: a new approach to incorporate patient's preferences into personalized pragmatic clinical trials. *Journal of Clinical Epidemiology*, *130*, 152-155.

Tueller, S., Ramirez, D., Cance, J. D., Ye, A., Wheeler, A. C., Fan, Z., Hornik, C. & Ridenour, T. A. (2022). Power analysis for idiographic (within-subject) clinical trials: Implications for treatments of rare conditions and precision medicine. *Behavior Research Methods*, 1-25.

Wittenborn, A. K., Liu, T., Ridenour, T. A., Lachmar, E. M., Mitchell, E. A., & Seedall, R. B. (2019). Randomized controlled trial of emotionally focused couple therapy compared to treatment as usual for depression: Outcomes and mechanisms of change. *Journal of Marital and Family Therapy*, *45*, 395-409.