Designing multi-center individually randomized group treatment trials

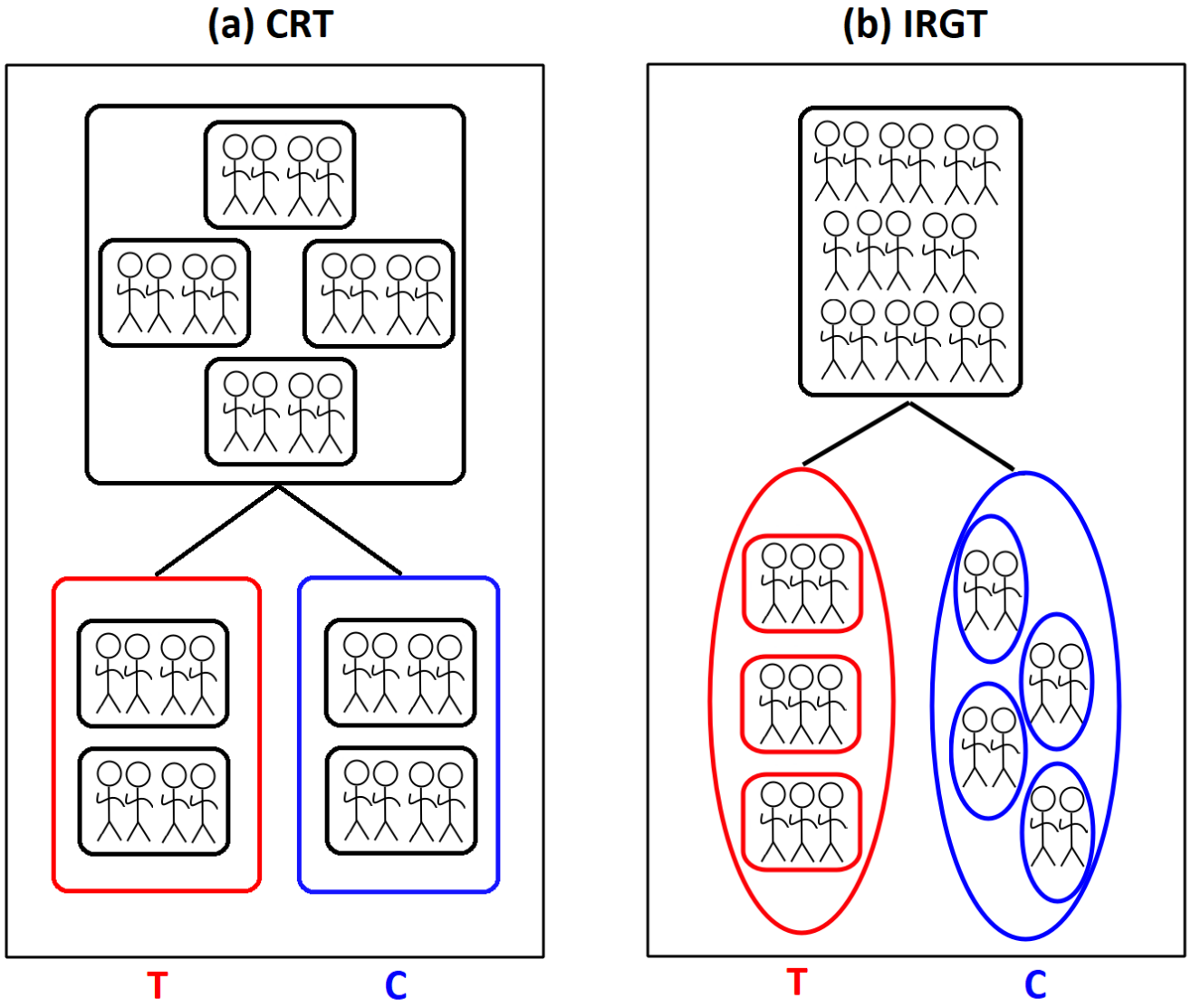
Guangyu Tong

Department of Internal Medicine (Cardiovascular Medicine)
Department of Biostatistics
Cardiovascular Medicine Analytics Center
Center for Methods in Implementation and Prevention Science
Yale University

Society of Clinical Trials 45th Annual Meeting
May 20-23, 2023

Two types of trials with correlated outcomes

(CRT: cluster randomized trial; IRGT: individually randomized group treatment trial)



Individually randomized group treatment trial (IRGT trial)

- IRGT: A type of trial design where individual participants are randomized to treatment conditions, but shared agents are needed during the treatment delivery
  - Each clinician may treat multiple patients
  - Therapy is delivered in a group setting
- Special case: IRGT design is also known as partially clustered design where active group-based treatment is compared with ungrouped standard care
- Clustering induced by treatment delivery has often been overlooked at both the design and analytical stages
- Ignoring the clustering may have led to over-optimistic findings

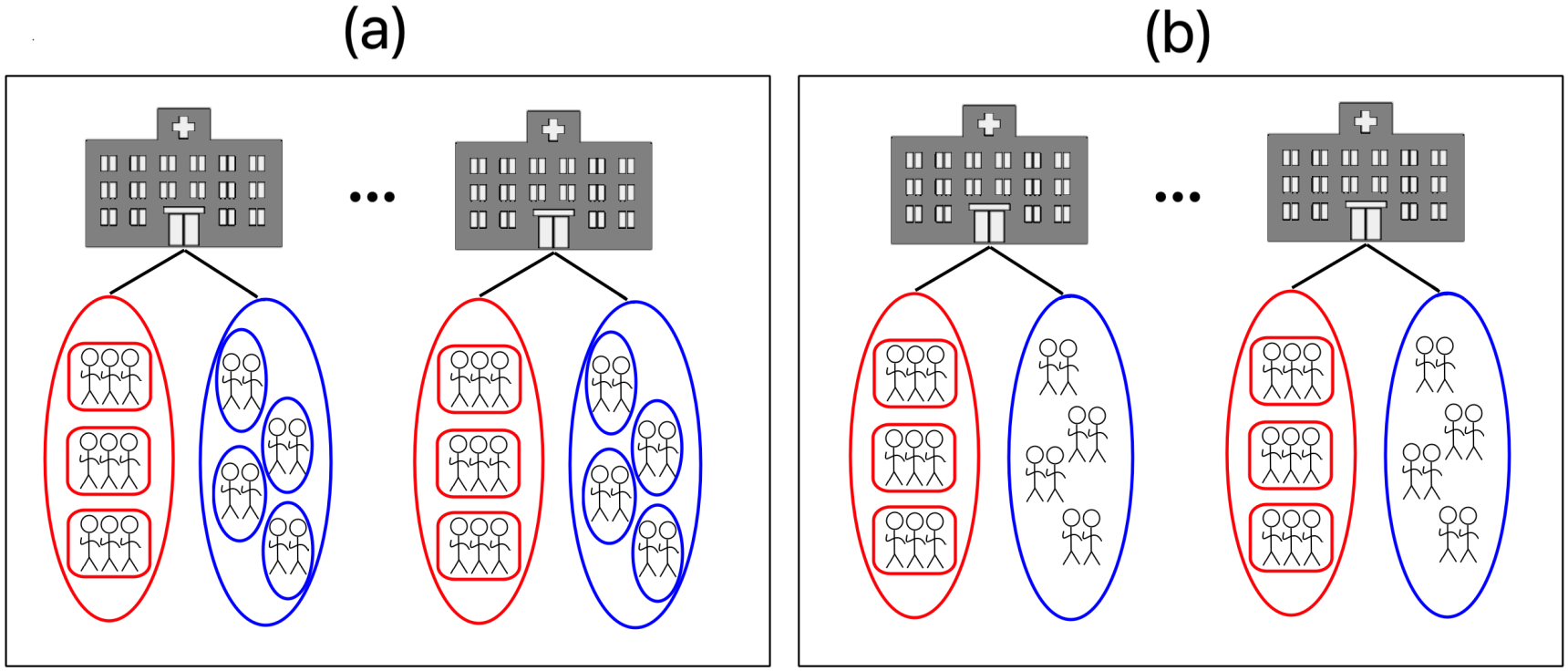
Design IRGTs to test the intervention effect

- Established literature on sample size consideration for testing the intervention effect (Roberts and Roberts 2005)
  - based on a linear mixed model with a treatment variable and a random cluster effect that allows between-arm heterogeneity
  - denote cluster sizes as m_1 and m_0 in treatment and control; ICC of the outcome as ρ_1 and $\rho_0 \Rightarrow$ variance inflation factor (VIF or design effect) due to cluster is VIF $= 1 + (m_x - 1)\rho_x$, where $x = 0, 1$
  - larger ICC leads to smaller power and larger required sample size
  - optimal allocation due to unequal design parameters between arms

Multi-center IRGT design

(Tong et al., 2024, Biometrical Journal)

- Because of limited treatment resources or participants at one location, an IRGT trial can be carried out across multiple centers.



Multi-center IRGT design

- In multi-center individually randomized trials, adjusting for the center effects in the analysis can improve the study power, especially when the between-center heterogeneity is substantial (Kahan and Morris, 2013)
- Appropriate methods that account for the within-center correlation structure are needed to help design multi-center IRGT trials and to elucidate the impact of ignoring the center effect in such studies

Linear mixed model with additional center effects

- A linear mixed model with additional sources of variability induced by centers (i: center; j: cluster; k: individual)

$$Y_{ijk} = \beta_0 + \beta_1 X_{ijk} + \lambda_i + X_{ijk}\theta_{ij1} + (1 - X_{ijk})\theta_{ij2} + X_{ijk}\epsilon_{ijk1} + (1 - X_{ijk})\epsilon_{ijk2}$$

  - center-effect: $\lambda_i \sim \mathcal{N}(0, \sigma^2_\lambda)$
  - arm-specific random effects: $\theta_{ij1} \sim \mathcal{N}(0, \sigma^2_{\theta_1})$ and $\theta_{ij2} \sim \mathcal{N}(0, \sigma^2_{\theta_2})$
  - arm-specific residual: $\epsilon_{ijk1} \sim \mathcal{N}(0, \sigma^2_{\epsilon_1})$ and $\epsilon_{ijk2} \sim \mathcal{N}(0, \sigma^2_{\epsilon_2})$

Five ICCs in the design of a multi-center IRGT trial

(Tong et al., 2024, Biometrical Journal)

Table 1 Expressions, definitions, and interpretations of ICC in a multi-center IRGT trial for within-therapist ICCs (ρ_1, ρ_2), between-therapist within-arm ICCs (α_1, α_2), and between-therapist between-arm ICC (κ). Variance parameters are defined as follows: $\sigma^2_\lambda = \text{Var}(\lambda_i)$ is center-level random-effect variance; $\sigma^2_{\theta_1} = \text{Var}(\theta_{ij1}), \sigma^2_{\theta_2} = \text{Var}(\theta_{ij2})$ are variances for arm-specific therapist-level random effects; $\sigma^2_{\epsilon_1} = \text{Var}(\epsilon_{ijk1}), \sigma^2_{\epsilon_2} = \text{Var}(\epsilon_{ijk2})$ are variances for arm-specific individual-level residual errors.

Expression	Definition	Interpretation
$\rho_1 = \dfrac{\sigma^2_\lambda + \sigma^2_{\theta_1}}{\sigma^2_\lambda + \sigma^2_{\theta_1} + \sigma^2_{\epsilon_1}}$	$\text{corr}(Y_{ijk}, Y_{ijk'} \mid X_{ijk} = X_{ijk'} = 1)$	correlation between two outcomes from two distinct individuals k, k' both in the treatment arm, treated by the same therapist j in the same center i
$\rho_2 = \dfrac{\sigma^2_\lambda + \sigma^2_{\theta_2}}{\sigma^2_\lambda + \sigma^2_{\theta_2} + \sigma^2_{\epsilon_2}}$	$\text{corr}(Y_{ijk}, Y_{ijk'} \mid X_{ijk} = X_{ijk'} = 0)$	correlation between two outcomes from two distinct individuals k, k' both in the control arm, treated by the same therapist j in the same center i
$\alpha_1 = \dfrac{\sigma^2_\lambda}{\sigma^2_\lambda + \sigma^2_{\theta_1} + \sigma^2_{\epsilon_1}}$	$\text{corr}(Y_{ijk}, Y_{ij'k'} \mid X_{ijk} = X_{ij'k'} = 1)$	correlation between two outcomes from two distinct individuals k, k' both in the treatment arm, treated by different therapists j, j' in the same center i
$\alpha_2 = \dfrac{\sigma^2_\lambda}{\sigma^2_\lambda + \sigma^2_{\theta_2} + \sigma^2_{\epsilon_2}}$	$\text{corr}(Y_{ijk}, Y_{ij'k'} \mid X_{ijk} = X_{ij'k'} = 0)$	correlation between two outcomes from two distinct individuals k, k' both in the control arm, treated by different therapists j, j' in the same center i
$\kappa = \dfrac{\sigma^2_\lambda}{\sqrt{\sigma^2_\lambda + \sigma^2_{\theta_1} + \sigma^2_{\epsilon_1}}\sqrt{\sigma^2_\lambda + \sigma^2_{\theta_2} + \sigma^2_{\epsilon_2}}}$	$\text{corr}(Y_{ijk}, Y_{ij'k'} \mid X_{ijk} = 0, X_{ij'k'} = 1)$	correlation between two outcomes from two distinct individuals k, k' in different treatment arms, treated by different therapists j, j' but the same center i

Implied constraints on the ICCs and variance parameters

(Tong et al., 2024, Biometrical Journal)

- α_1/ρ_1 and α_2/ρ_2 are bounded between 0 and 1
- The between-therapist between-arm ICC is then induced as $\kappa = \sqrt{\alpha_1\alpha_2}$
- If we define the total variance of the treatment arm and the control arm as $\sigma_1^2 = \sigma_\lambda^2 + \sigma_{\theta_1}^2 + \sigma_{\epsilon_1}^2$, and $\sigma_2^2 = \sigma_\lambda^2 + \sigma_{\theta_2}^2 + \sigma_{\epsilon_2}^2$, then we have the constraint of $\alpha_2\sigma_2^2 = \alpha_1\sigma_1^2$
- In partially nested multi-center IRGT, we have $\rho_2 = \alpha_2$

Doubly nested exchangeable correlation structure

- All individual-level outcomes at each center follow the following correlation structure

$$\boldsymbol{S}_i = \begin{pmatrix} \boldsymbol{R_1} & \kappa \boldsymbol{J}_{T_1 N_1 \times T_2 N_2} \\ \kappa \boldsymbol{J}_{T_2 N_2 \times T_1 N_1} & \boldsymbol{R_2} \end{pmatrix}$$

  - The correlation matrices for each arm are

$$\boldsymbol{R}_1 = (1-\rho_1)\boldsymbol{I}_{T_1 N_1} + (\rho_1 - \alpha_1)\boldsymbol{I}_{T_1} \otimes \boldsymbol{J}_{N_1} + \alpha_1 \boldsymbol{J}_{T_1 N_1},$$
$$\boldsymbol{R}_2 = (1-\rho_2)\boldsymbol{I}_{T_2 N_2} + (\rho_2 - \alpha_2)\boldsymbol{I}_{T_2} \otimes \boldsymbol{J}_{N_2} + \alpha_2 \boldsymbol{J}_{T_2 N_2}.$$

  - Each arm has a nested exchangeable correlation structure (Li et al. 2019)
  - T_1, T_2 are the number of therapists (clusters) in each arm; N_1, N_2 are the cluster sizes
- The variance of β_1 can be derived under generalized least squares

Variance formula of the intervention effect

- Variance formula expressed in variance

$$\mathrm{var}(\widehat{\beta}_1) = \frac{N_1\sigma^2_{\theta_1} + \sigma^2_{\epsilon_1}}{IT_1N_1} + \frac{N_2\sigma^2_{\theta_2} + \sigma^2_{\epsilon_2}}{IT_2N_2}$$

- Variance formula expressed in ICCs

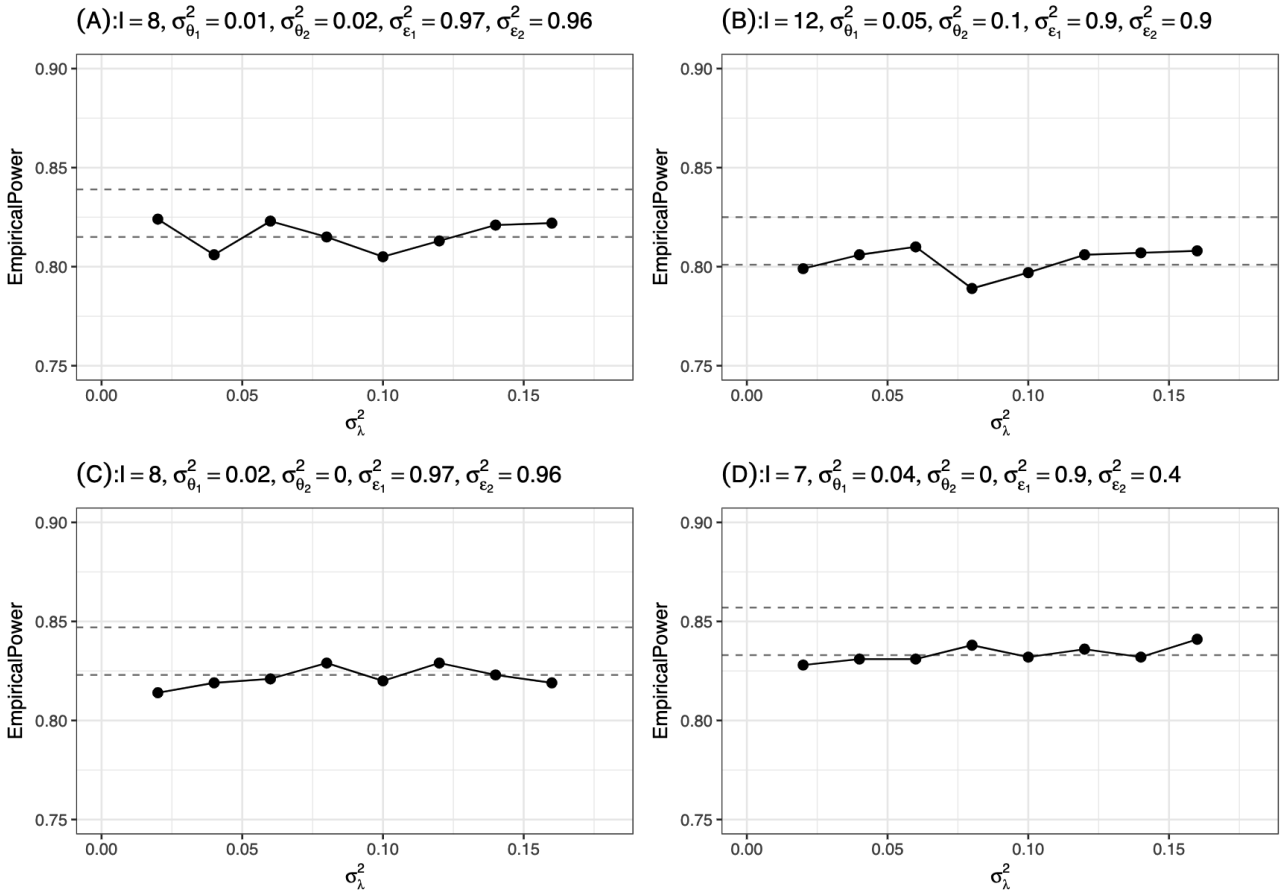
$$\mathrm{var}(\widehat{\beta}_1) = \frac{\sigma_1^2\{1 + (N_1 - 1)\rho_1 + N_1(T_1 - 1)\alpha_1\}}{IT_1N_1} + \frac{\sigma_2^2\{1 + (N_2 - 1)\rho_2 + N_2(T_2 - 1)\alpha_2\}}{IT_2N_2} - \frac{2\sigma_1\sigma_2\sqrt{\alpha_1\alpha_2}}{I}$$

  - ρ_1 and ρ_2 are the within-therapist ICCs
  - α_1 and α_2 are the between-therapist, within-arm ICCs

- The sample size and power formula can then be derived under a two-sided z-test

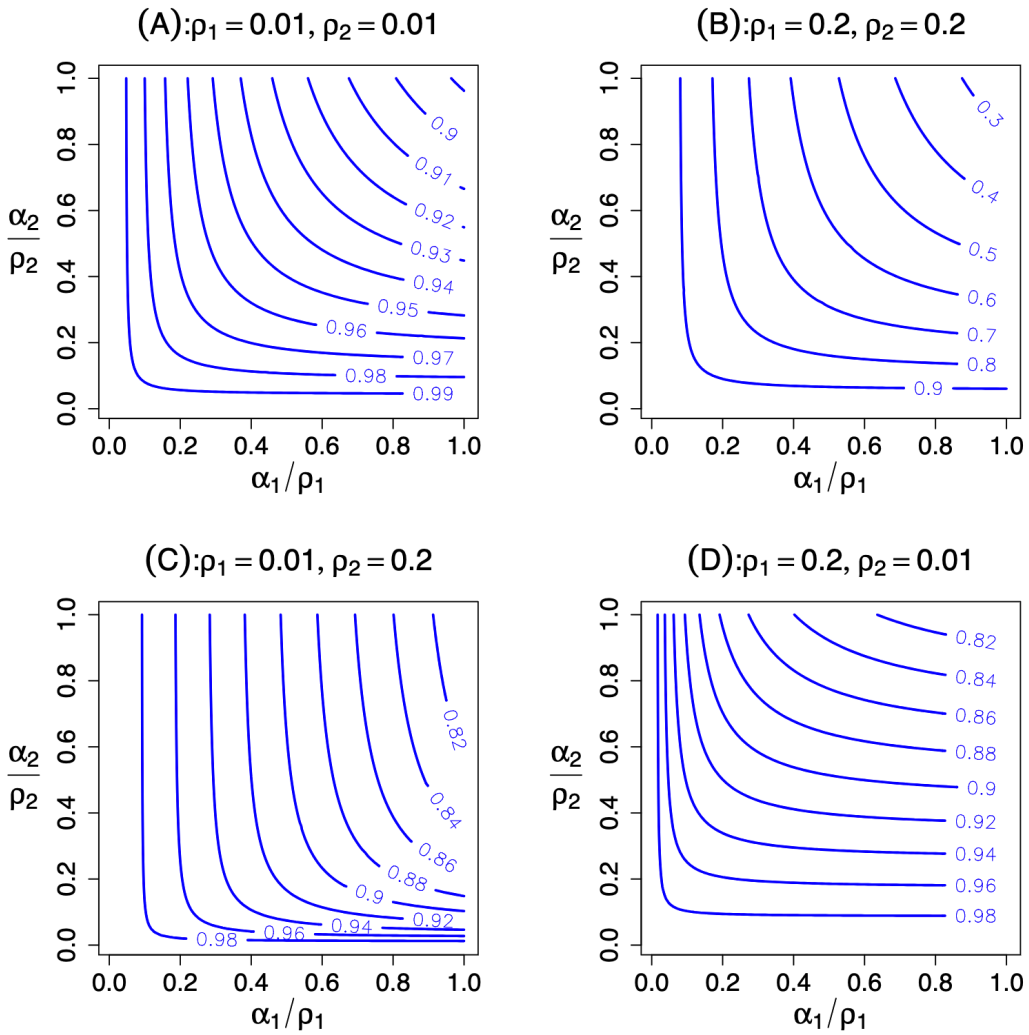
The impact of σ_λ^2 on power when using variance in design

Figure 3 Empirical power of the treatment effect when varying σ_λ^2. The number of therapists per center for two arms is $T_1 = 4, T_2 = 6$ respectively for all panels. The cluster sizes are $N_1 = 15$ and $N_2 = 10$. The required number of sites I for each panel is fixed, which is labeled in the panel title. Two dashed lines for each panel correspond to a nominal scope of empirical power.



The impact of ICCs on variance reduction

Figure 2 The ratio of the variance of treatment effect estimator accounting for the center effect to the variance not accounting for center effect varies as a function of α_1/ρ_1 and α_2/ρ_2. The number of therapists per site for each arm is $T_1 = 4, T_2 = 6$ respectively. The number of individuals per therapist for each arm is $N_1 = 15, N_2 = 10$ respectively. Each panel corresponds to one combination of ICCs which is labeled in the panel title.



Optimal allocation and partially nested design

- Optimal allocation ratio on therapists to reach the maximized power is

$$R^{opt} = \frac{T_1^{opt}}{T_2^{opt}} = \frac{\sigma_1\sqrt{N_2\{1+(N_1-1)\rho_1 - N_1\alpha_1\}}}{\sigma_2\sqrt{N_1\{1+(N_2-1)\rho_2 - N_2\alpha_2\}}}$$

  - To ensure this optimal allocation ratio is valid, we need to have $1+(N_1-1)\rho_1 - N_1\alpha_1 > 0$ and $1+(N_2-1)\rho_2 - N_2\alpha_2 > 0$

- For partially nested IRGT design, the optimal allocation ratio can be obtained by setting the cluster size of the arm with no clustering to 1

Optimal allocation accounting for cost

- Assume C_0 is the overhead cost of the trial; B is the cost of each center; c_1 and c_2 are the cost per therapist in each arm; s_1 and s_2 are the cost per participant per therapist in each arm. The total cost or grand budget of a multi-center IRGT trial is then

$$C = C_0 + I(B + T_1(c_1 + N_1 s_1) + T_2(c_2 + N_2 s_2))$$

- Optimal allocation ratio can be obtained, and then the number of therapists per center for the treatment arm can be derived as

$$T_1^{opt} = \frac{(C - C^*)/I}{\sqrt{c_1^* c_2^*} \times \dfrac{N_1 \sigma_2^2 \{1 + (N_2 - 1)\rho_2 - N_2 \alpha_2\}}{N_2 \sigma_1^2 \{1 + (N_1 - 1)\rho_1 - N_1 \alpha_1\}} + c_1^*},$$

where $c_1^* = c_1 + N_1 s_1$, $c_2^* = c_2 + N_2 s_2$, and $C^* = C_0 + IB$

Simulation design

- Simulation parameters for multi-center IRGT (240 scenarios)
  - Total variance of the treatment arm as $\sigma_1^2 = 1$
  - Within-therapist ICCs are $\rho_1, \rho_2 \in \{0.01, 0.05, 0.10, 0.20\}$
  - Between-therapist ICCs as fractions of within-therapist ICCs, $\alpha_1 = \gamma\rho_1$ and $\alpha_2 = \gamma'\rho_2$, where $\gamma, \gamma' \in \{0, 0.2, 0.5, 0.8, 1.0\}$
  - Number of therapists per center for each arm as $(T_1, T_2) \in \{(2, 4), (4, 6), (5, 10)\}$
  - The corresponding number of patients per therapist for each arm as $(N_1, N_2) \in \{(30, 15), (15, 10), (12, 6)\}$
- Simulation parameters for multi-center partially nested design (32 scenarios)
  - Within-therapist ICC $\rho_1 \in \{0.02, 0.04, 0.08, 0.20\}$
  - $\alpha_1 = 0.25\rho_1$, and $\rho_2 = 0.5\rho_1$ or $\rho_2 = 0.75\rho_1$
  - Number of therapists per center as $T_1 \in \{2, 4, 6, 10\}$
  - Number of patients per therapist as $N_1 \in \{30, 15, 10, 6\}$

Simulation design

- All sets of simulations follow four steps
  - Estimate the required sample size based on the variance formula. Here, the number of centers is rounded up to the closest integer
  - Simulate outcomes from the linear mixed model under the null ($\beta_1 = 0$) or the alternative ($\beta_1 = 0.2$) separately
  - Fit linear mixed models by using the "lme" function from the "nlme" package in R with doubly nested exchangeable correlation structure and restricted maximum likelihood
  - Obtain the p value associated with testing the treatment effect, β_1, based on the two-sided z-test.
- Each scenario repeats the above steps for $\boldsymbol{B} = 5000$ times
- Type I error rate and power are average over results obtained in $\boldsymbol{B}$ times

Simulation results

Table 2 Simulation results for planning two-arm multi-center IRGT trials with individuals nested in therapists in both arms under 80% power, two-sided 5% type I error rate, and $\gamma = \{0.2, 0.5, 0.8\}$. Results include the required number of sites (I), empirical type I error rate (ψ), predicted power (ϕ_{pre}), and empirical power (ϕ_{emp}). The numbers of therapists and the numbers of patients per center for two arms are $T_1 = 4$, $T_2 = 6$, $N_1 = 15$, and $N_2 = 10$.

		$\alpha_1 = 0.2\rho_1, \alpha_2 = 0.2\rho_2$				$\alpha_1 = 0.5\rho_1, \alpha_2 = 0.5\rho_2$				$\alpha_1 = 0.8\rho_1, \alpha_2 = 0.8\rho_2$			
ρ_1	ρ_2	I	ψ	ϕ_{pre}	ϕ_{emp}	I	ψ	ϕ_{pre}	ϕ_{emp}	I	ψ	ϕ_{pre}	ϕ_{emp}
0.01	0.01	8	0.045	0.843	0.838	7	0.042	0.807	0.792	7	0.039	0.820	0.799
	0.05	5	0.044	0.839	0.820	5	0.042	0.858	0.837	5	0.043	0.878	0.854
	0.10	5	0.044	0.865	0.841	4	0.040	0.806	0.767	4	0.039	0.830	0.791
	0.20	5	0.044	0.878	0.852	4	0.040	0.822	0.785	4	0.042	0.846	0.810
0.05	0.01	23	0.049	0.807	0.812	22	0.050	0.812	0.815	21	0.048	0.817	0.807
	0.05	10	0.054	0.820	0.823	9	0.048	0.833	0.823	8	0.044	0.848	0.844
	0.10	8	0.045	0.807	0.795	7	0.048	0.821	0.817	6	0.045	0.840	0.825
	0.20	8	0.044	0.848	0.837	6	0.043	0.813	0.795	5	0.043	0.833	0.813
0.10	0.01	42	0.055	0.801	0.802	40	0.049	0.806	0.805	37	0.051	0.800	0.807
	0.05	16	0.048	0.807	0.817	14	0.055	0.821	0.827	11	0.049	0.807	0.805
	0.10	13	0.054	0.817	0.818	10	0.054	0.801	0.802	8	0.046	0.824	0.819
	0.20	11	0.047	0.807	0.791	9	0.048	0.829	0.814	6	0.045	0.807	0.792
0.20	0.01	81	0.048	0.803	0.811	76	0.050	0.802	0.805	71	0.046	0.802	0.804
	0.05	29	0.052	0.812	0.819	24	0.058	0.813	0.819	19	0.051	0.814	0.820
	0.10	22	0.050	0.807	0.808	17	0.050	0.807	0.816	12	0.053	0.807	0.810
	0.20	19	0.049	0.814	0.813	14	0.052	0.817	0.809	9	0.050	0.822	0.811

Data example 1: Schema Therapy for Personality Disorders (STPD) trial

(Bamelis et al., 2014)

- 323 personality disorders patients from 12 Dutch mental health institutes, and participants were individually randomized to one of the three arms, schema therapy (sample size is 147) or treat-as-usual (sample size is 135), and clarification-oriented psychotherapy (sample size is 41)
- The schema therapy arm has 66 therapists with an average cluster size of 2.2, whereas the treat-as-usual arm has 79 therapists with an average cluster size of 1.7
- A key secondary outcome-the global assessment of functioning (GAF) scale score (0-100) measured at the 36 month, a scale of 0 to 100 with a standard deviation usually around 13

Data example 2: Comparison of Outcomes of Antibiotic Drugs and Appendectomy (CODA) trial

(CODA-Collaborative, 2020)

- CODA compared antibiotic therapy to surgery as an alternative treatment (1:1 ratio) for appendicitis with a total of 1552 adults recruited from 25 centers in the United States
- Surgical treatments were delivered by a total of 48 surgeons across centers, whereas the antibiotic treatment was ungrouped. The average cluster size is 16
- The primary outcome is quality of life (measured by the EuroQol at 4 week), a continuous score ranging from 0 to 1, with the standard deviation estimated as 0.13
- The original sample size calculation expected an effect size of 0.008

Key considerations for multi-center IRGT design

- If the variance components of therapist-level random effects are considered as input parameters in the design stage, accounting for the additional center-level variance component has no impact on the sample size estimation
- If the ICC parameters are considered, accounting for the additional center-level correlation can lead to sample size reduction, and the magnitude of reduction depends on the amount of between-therapist correlation
- Optimal allocation (with options to accommodate costs) is recommended to get a more efficient design

Key considerations for multi-center IRGT design

- One needs to be aware of the constraints on the ICC and variance parameters at the design
  - If at least one of the arms has high therapist ICC (ρ_1, ρ_2), it is potentially important to consider the center effect
  - If the total variance of the two arms is drastically different, then it is potentially important to consider the center effect
- When the number of centers is small (e.g., 7 or below), the sample size result using our method can be slightly under power
- Please find more content in, **Tong G.**, Tong J., Li F. (2024). Designing multi-center individually randomized group treatment trials. *Biometrical Journal*, 66(1), 1-19.

Acknowledgement and funding

- Coauthors: This is joint work Jiaqi Tong (PhD student at Yale Biostatistics) and Fan Li (Professor at Yale Biostatistics)
- The authors thank Dr. Patrick Heagerty (UWash Biostatistics) and Dr. Denise Esserman (Yale Biostatistics) for pointing us to relevant applications and helpful discussions.
- Funding support: This work was partially supported by the Patient-Centered Outcomes Research Institute Award(PCORI® Award ME-2020C3-21072, PI: Li) titled, "Developing New Methods for Assessing Heterogeneity of Treatment Effect in Cluster Randomized Trials."

References

- Roberts C. and Roberts SA. (2005). Design and analysis of clinical trials with clustering effects due to treatment. *Clinical Trials*, 2(2): 152-162.
- Li F., Forbes AB., Turner EL. and Preisser JS. (2019) Power and sample size requirements for GEE analyses of cluster randomized crossover trials. *Statistics in Medicine*, 38(4): 636-649.
- CODA-Collaborative (2020) A randomized trial comparing antibiotics with appendectomy for appendicitis. *New England Journal of Medicine*, 383(20): 1907â1919.
- Bamelis LL, Evers SM, Spinhoven P and Arntz A (2014) Results of a multicenter randomized controlled trial of the clinical effectiveness of schema therapy for personality disorders. *American Journal of Psychiatry*, 171(3): 305-332.