Essential Ingredients and Innovations in the Design and Analysis of Group-Randomized Trials

13th Annual Conference on Statistical Issues in Clinical Trials

April 12, 2021

David Murray, Ph.D.,
Associate Director for Prevention,
Director, Office of Disease Prevention
National Institutes of Health

National Institutes of Health
Office of Disease Prevention

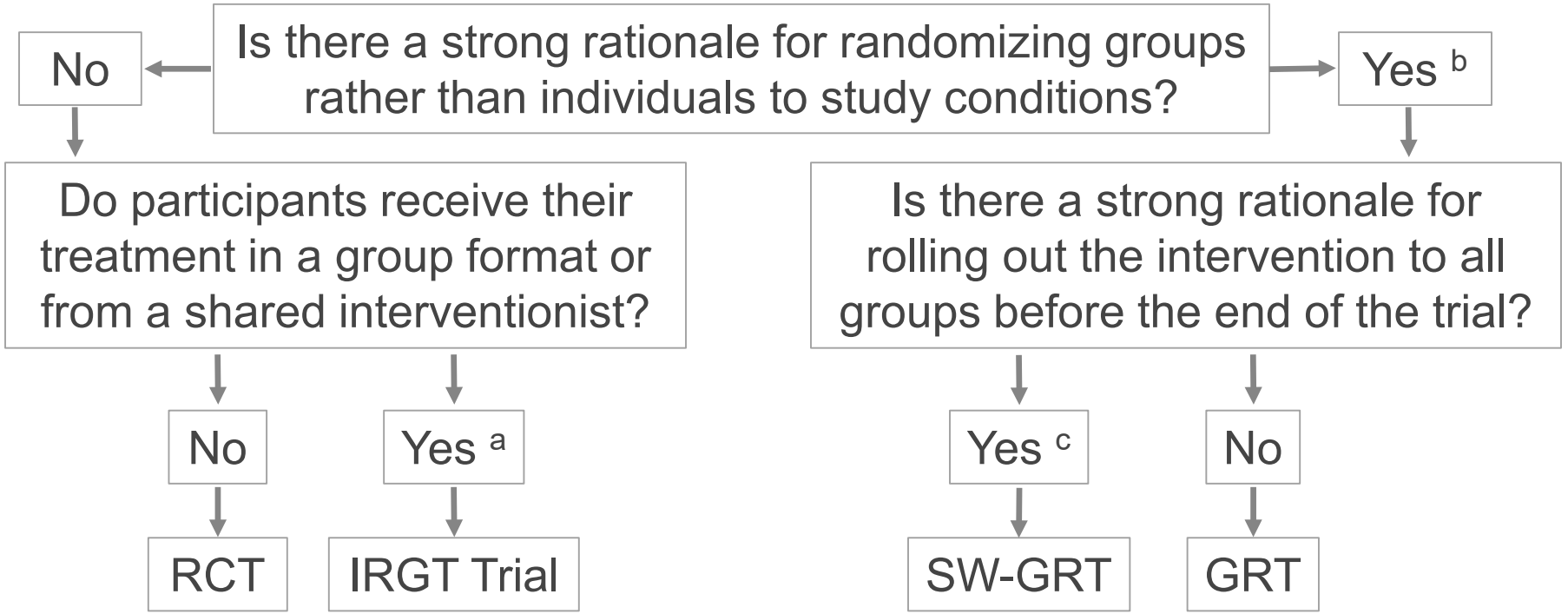
Three Kinds of Randomized Trials

- Randomized Clinical Trials (RCTs)
  - Individuals randomized to study conditions with no interaction among participants after randomization (no group sessions, virtual interaction, or shared intervention agent)
    - Most drug trials
- Individually Randomized Group Treatment Trials (IRGTs)
  - Individuals randomized to study conditions with interaction among participants after randomization or with a shared intervention agent
    - Many surgical trials
    - Many behavioral trials
- Group-Randomized Trials (GRTs)
  - Groups randomized to study conditions with interaction among the members of the same group before and after randomization
    - Many trials conducted in communities, worksites, schools, clinics, etc.

Two Kinds of Group-Randomized Trials

- Parallel GRT
  - Separate but parallel intervention and control conditions throughout the trial, with no crossover.
- Stepped Wedge GRT
  - All groups start in the control condition.
  - All groups crossover to the intervention condition, but in a random order and on a staggered schedule.
  - All groups receive the intervention before the end of the study.

Choosing Among These Designs



[a] If the intervention is delivered through a physical or a virtual group, or through shared interventionists who each work with multiple participants, positive ICC can develop over the course of the trial.

[b] There may be logistical reasons to randomize groups or it may not be possible to deliver the intervention to individuals without substantial risk of contamination.

[c] There may be good political or logistical reasons to roll out the intervention to all groups before the end of the trial.

Adapted from Murray DM, Taljaard M, Turner EL, George SM. Essential Ingredients and Innovations in the Design and Analysis of Group-Randomized Trials. Annual Review of Public Health. 2020;41:1-19. PMID31869281.

Analysis Issues in a GRT or IRGT

- Nested factors must be modeled as random effects (Zucker, 1990).
- The variance of any group-level statistic will be larger.
- The df to estimate the group-level component of variance will be based on the number of groups and is often limited.
  - This is almost always true in a GRT and can be true in an IRGT.
- Any analysis that ignores the extra variation or the limited df will have a Type I error rate that is inflated, often badly (Cornfield, 1978).
  - Type I error rate may be 30-50% in a GRT, even with small ICC
  - Type I error rate may be 15-25% in an IRGT, even with small ICC
- Extra variation and limited df always reduce power.

Zucker DM. An analysis of variance pitfall: The fixed effects analysis in a nested design. Educ and Psych Measurement. 1990;50(4):731-8.

Cornfield J. Randomization by group: a formal analysis. American Journal of Epidemiology. 1978;108(2):100-2.

Analysis Issues for SW-GRTs

- Crossing of groups with study conditions often reduces the impact of the ICC compared to a parallel GRT, either improving power or allowing a smaller study.
- There are other potential sources of bias in the SW-GRT:
  - The intervention is confounded with time.
  - The intervention effect may vary over time.
  - The intervention effect may vary by group.
  - Patterns of correlation may vary over time.
- Any analysis that assumes that the intervention effect is constant over time and across groups, and that the pattern of correlation is constant, may be biased.
- Compared to a parallel GRT, SW-GRTs are at greater risk to the effects of external events that affect the outcomes of the trial.

Evolution of the Methods

- The co-authors identified 4514 candidate reports through 2018.
- Preliminary screening identified 926 focused on GRTs, IRGTs, or SW-GRTs.
- The Relative Citation Ratio (RCR) and citation counts were used to identify influential reports.
  - The RCR is an article-level and field-independent metric that reflects the degree to which an article is cited relative to the articles that appear alongside it in reference lists of other papers – the co-citation network (Hutchins et al., 2016).
    - RCR values were available for 85.4% of the reports.
    - GRTs – RCR in top 1% or >200 citations
    - IRGTs – RCR in top 5% or >100 citations
    - SW-GRTs – RCR in top 2.5% or >150 citations
- 50 influential reports were identified (Murray et al., 2020).

Hutchins BI, Yuan X, Anderson JM, Santangelo GM. Relative Citation Ratio (RCR): A New Metric That Uses Citation Rates to Measure Influence at the Article Level. PLoS Biol. 2016;14(9):e1002541. PMID27599104.

Murray DM, Taljaard M, Turner EL, George SM. Essential Ingredients and Innovations in the Design and Analysis of Group-Randomized Trials. Annual Review of Public Health. 2020;41:1-19. PMID31869281.

Group-Randomized Trials

- Cornfield (1978) – the two penalties of extra variation and limited df
- Donner et al. (1981) – methods for sample size and analysis
- Feldman (1988) – random coefficients models
- Murray et al. (1990) – design, analytic, and sample size methods
- Rao & Scott (1992) – methods for analysis of binary outcomes
- Donner & Klar (1992) – methods for meta-analysis
- Hedeker & Gibbons (1994) – methods for ordinal outcomes
- Hedeker & Gibbons (1996) – software for ordinal outcomes
- Bland et al. (1997) – brief summary of the design and analytic issues
- Raudenbush (1997) – analytic methods and optimal design
- Murray (1998) – first textbook on design and analytic methods
- Guiliford et al. (1999) – estimates of intraclass correlation

Group-Randomized Trials

- Donner & Klar (2000) – second textbook on design and analytic methods
- Krull & MacKinnon (2001) – methods for mediation analysis
- Eccles et al. (2003) – alternative research designs, including GRTs
- Adams et al. (2004) – estimates of intraclass correlation
- Murray et al., (2004) – state of the practice for design and analysis of GRTs
- Campbell et al. (2004, 2012) – the CONSORT statement for GRTs
- Hedges & Hedberg (2007) – estimates of intraclass correlation
- Hayes & Moulton (2009) – third textbook on design and analytic methods
- Emsley et al. (2010) – methods for mediation and moderation
- Eldridge & Kerry (2012) – fourth textbook on design and analytic methods
- Zou & Donner (2013) – modified Poisson regression model
- Grant et al. (2013) – methods for process evaluation

Individually Randomized Group-Treatment Trials

- Whiting-O'Keefe et al. (1984) – the correct unit of analysis in medical experiments
- Crits-Cristoph & Mintz (2001) – implications of therapist effects
- Nye et al. (2004) – estimates of variance components in IRGTs
- Baldwin et al. (2005) – the impact of ignoring therapist effects
- Roberts & Roberts (2005) – analytic methods for IRGTs
- Boutron et al. (2008) – CONSORT statement for non-pharmacologic interventions
- Kahan & Morris (2013) – sources of clustering in individually randomized trials
- Heo et al. (2017) – sample size methods for IRGTs
- Sterba (2017) – modeling developments for partially nested designs

Stepped Wedge Group-Randomized Trials

- Hussey & Hughes (2005) – analytic methods for SW-GRTs
- Brown & Lilford (2006) – state of the practice for design and analysis of SW-GRTs
- Mdege et al. (2011) – state of the practice for design and analysis of SW-GRTs
- Woertman et al. (2013) – sample size methods for SW-GRTs
- Baio et al. (2015) – sample size methods for SW-GRTs
- Beard et al. (2015) – state of the practice for design and analysis of SW-GRTs
- Copas et al. (2015) – typology for SW-GRT designs
- Davey et al. (2015) – state of the practice for design and analysis of SW-GRTs
- Hemming et al. (2015) – rationale, design, and analysis for SW-GRTs
- Hemming et al. (2015) – sample size methods for SW-GRTs
- Girling & Hemming (2016) – statistical efficiency and optimal design for SW-GRTs
- Hemming & Taljaard (2016) – sample size methods for SW-GRTs
- Scott et al. (2017) – small sample correction for GEE in SW-GRTs

Limitations

- The review considered papers published through 2018.
- Papers published in 2017 and 2018 will usually have fewer citations and may not yet have had an RCR value in 2019.
- Taken together, these issues may underestimate the influence of papers published in 2017 and 2018.
- Papers published after 2018 were not considered.

NIH Resources

- Pragmatic and Group-Randomized Trials in Public Health and Medicine
  - https://prevention.nih.gov/grt
  - 7-part online course on GRTs and IRGTs
- Mind the Gap Webinars
  - https://prevention.nih.gov/education-training/methods-mind-gap
    - SW-GRTs for Disease Prevention Research (Monica Taljaard, July 11, 2018)
    - Design and Analysis of IRGTs in Public Health (Sherri Pals, April 24, 2017)
    - Research Methods Resources for Clinical Trials Involving Groups or Clusters (David Murray, December 13, 2017)
- Research Methods Resources Website
  - https://researchmethodsresources.nih.gov/
  - Material on GRTs and IRGTs and a sample size calculators for GRTs and IRGTs.