Unplanned Disruptions in Pragmatic Clinical Trials: Examples from the Health Care Systems Collaboratory

September 1, 2020

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

Pragmatic Trials

- Pragmatic and explanatory trials were first described by Schwartz & Lellouch (1967).
  - Explanatory trials test causal research hypotheses.
  - Pragmatic trials help users choose between options for care.
- Similar to efficacy and effectiveness trials (Cochrane, 1971).
  - Efficacy trials evaluate an intervention under controlled conditions.
  - Effectiveness trials evaluate an intervention under real-world conditions.

- Schwartz, D., & Lellouch, J. Explanatory and pragmatic attitudes in therapeutical trials. Journal of Chronic Diseases, 1967, 20(8), 637-648. PMID4860352
- Cochrane, A.L. Effectiveness and efficacy: random reflections on health services. Nuffield Provincial Hospitals Trust, London, 1971. (cited in Flay, Brian R. Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. Preventive Medicine, 1986, 15(5), 451-474. PMID3534875.

Designs Commonly Used in Pragmatic Trials

- Randomized Controlled Trial (RCT)
  - Individuals randomized to study conditions with no interaction among participants after randomization (no group sessions, virtual interaction, or shared interventionist).
- Individually Randomized Group Treatment Trial (IRGT)
  - Individuals randomized to study conditions with intervention-induced correlation among observations taken from participants who receive all or part of their intervention in a group format or through a shared interventionist (live or virtual).
- Parallel Group-Randomized Trial (GRT)
  - Groups randomized to study conditions with correlation among the members of the same group before and after randomization.
- Stepped-Wedge GRT (SW-GRT)
  - Correlation among the members of the same group before and after intervention.
  - All groups start in the control condition and crossover to the intervention condition in a random order and on a staggered schedule.

Impact on the Design

- Randomized clinical trials
  - There is usually good opportunity for randomization to distribute potential confounders evenly, as most RCTS have N>100.
- Individually randomized group treatment trials
  - There may be less opportunity for randomization to distribute potential confounders evenly, as many IRGTs have N<100.
- Parallel group-randomized trials
  - In any single realization, there is limited opportunity for randomization to distribute all potential confounders evenly as often G<50.
- Stepped wedge GRTs
  - Observing each group under both study conditions avoids most confounding.
  - However, intervention effects are confounded with calendar time by design.
  - SW-GRTs are inherently less rigorous than parallel GRTs and should be considered only when a parallel GRT is not appropriate.

Impact on the Analysis 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, 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.
  - 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.

Impact on the Analysis 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 by design.
  - The intervention effect may vary over time.
  - The intervention effect may vary by group.
  - Patterns of correlation may vary over time.
- Any analysis that ignores these potential sources may be biased.
- Compared to a parallel GRT, SW-GRTs are at greater risk to the effects of external events that can affect the outcomes of the trial.

- Hemming K, Haines TP, Chilton PJ, Girling AJ, Lilford RJ. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. BMJ. 2015;350:h391. PMID: 25662947.
- Li F, Hughes JP, Hemming K, Taljaard M, Melnick ER, Heagerty PJ. Mixed-effects models for the design and analysis of stepped wedge cluster randomized trials: An overview. Stat Methods Med Res. In press.

The Warning

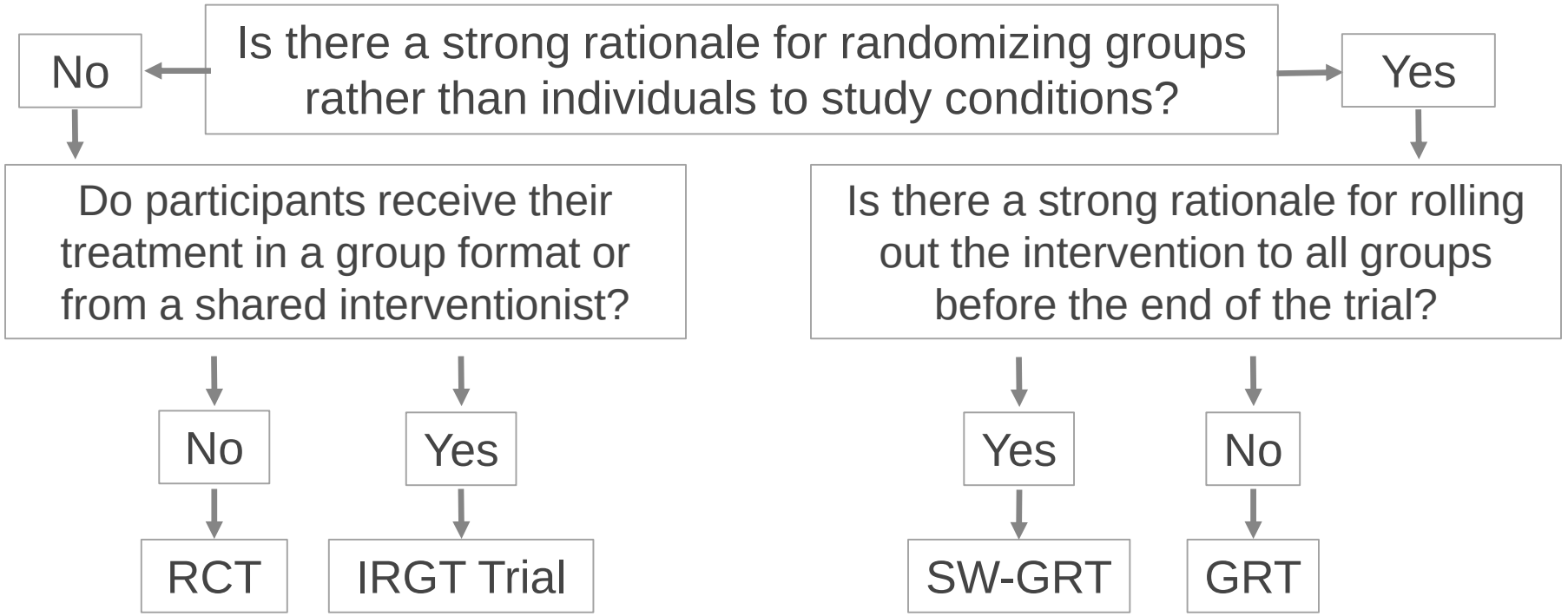
Randomization by cluster accompanied by an analysis appropriate to randomization by individual is an exercise in self-deception, however, and should be discouraged.

Cornfield (1978)

- Though Cornfield's remarks were addressed only to GRTs, they also apply to IRGTs, and to SW-GRTs.

- Cornfield J. Randomization by group: a formal analysis. Am J Epi. 1978;108(2):100-2.

Choosing Among These Designs



- Murray DM, Taljaard M, Turner EL, George SM. Essential Ingredients and Innovations in the Design and Analysis of Group-Randomized Trials. Annu Rev Public Health. 2020;41:1-19. PMID: 31869281.

The Need for GRTs, IRGTs, and SW-GRTs

- An RCT is the best comparative design when individual randomization is possible without post-randomization interaction.
- An IRGT is the best comparative design whenever...
  - Individual randomization is possible but there are good reasons to deliver the intervention in a group format or through a shared interventionist
- A GRT is the best comparative design whenever the investigator wants to evaluate an intervention that…
  - Manipulates the social or physical environment or cannot be delivered to individuals without risk of contamination
- An SW-GRT is an alternative to a parallel GRT if…
  - Preliminary evidence makes it unethical to withhold the intervention.
  - It is impossible to implement the intervention in all groups simultaneously.
  - External events are unlikely to affect the outcomes before the end of the trial.

Health Care Systems Collaboratory

- Supported by the NIH Common Fund beginning in 2012; expanded to coordinate the Pain Management Collaboratory studies in 2019
- Engages health care delivery organizations as research partners in the conduct of pragmatic clinical trials.
- Generates and disseminates best practices for conducting pragmatic trials.
- 1 Coordinating Center funded using a U54 award.
- 19 clinical trials funded using UG3/UH3 awards that allow investigators one year to refine their plans and up to five years to conduct the trial.

	RCT	GRT	SW-GRT	IRGT	Total
Completed	1	6	2	0	9
Underway	1	3	2	0	6
Planning	0	1	1	2	4
Total	2	10	5	2	19

Role of the Biostatistics and Design Working Group

- Comprised of methodologists from NIH, the Coordinating Center, and each of the participating trials.
- The WG reviews each project during the first part of their UG3 year to consider their design, analytic plan, and sample size calculations, and to offer advice to help them submit as strong an application as possible for the UH3 project.
- The WG continues to interact with trials funded for the UH3 as they encounter unexpected methodological issues.
- The COVID pandemic presented itself as just such an issue earlier this year.
- The WG has been working with the current trials to help them think through the challenges posed by COVID and to make adaptations.
- This discussion is guided by the PICOT evaluation framework and by the CONSORT reporting guidelines for the various designs.
  - PICOT stands for population, intervention, comparator, outcome, time frame
  - Haynes et al. (2006) *Clinical epidemiology: how to do clinical practice research*.

Examples from the Collaboratory

- Stepped Wedge Group-Randomized Trial underway when the pandemic began
  - Promoting Effective Advance Care Planning in the Elderly
- Parallel Group-Randomized Trials underway when the pandemic began
  - HiLo
  - Guiding Good Choices for Health
- Individually Randomized Group Treatment Trial in planning when the pandemic began
  - Pragmatic Trial of Acupuncture for Chronic Low Back Pain in Older Adults

Promoting Effective Advance Care Planning in the Elderly (ACP-PEACE)

- Original design: stepped wedge group-randomized trial

Clinics Per Site	Step						
	Baseline	1	2	3	4	5	6
1, 2							
3, 4							
5, 6							
7, 8							
9, 10							
11, 12							

- Groups: 36 oncology clinics in 3 Sites, 4500 cancer patients >65.
- Intervention: targets physicians (skills training) and patients (videos)
- Outcomes: ACP documentation (yes vs. no)
- Analysis: mixed-model logistic regression with clinic as a nested random effect.

ACP-PEACE Redesigned

- COVID struck during Step 2 and all clinics increased advanced care planning.
- This affected the primary outcome and interrupted training for Step 2.
- Solution: use Step 2 as the baseline in the remaining clinics.
  - Allows test of the intervention pre COVID in 6 clinics.
  - Allows test of the intervention post COVID in 30 clinics.
  - Allows assessment of effect of COVID on ACP rates in 30 clinics using control data.

Clinics Per Site	Step						
	Baseline	1	2	3	4	5	6
1, 2							
			Baseline	1	2	3	4
3, 4, 5							
6, 7							
8, 9, 10							
11, 12							

HiLo

- Design: multi-center, open label, parallel group-randomized non-inferiority trial
- Groups: 120 dialysis facilities with 4400 patients (g=60, m=37)
- Intervention: two phosphate management strategies guided by facility dieticians
- Outcomes: composite of mortality and hospitalization outcomes
- Analysis: permutation test or mixed model

HiLo

- COVID struck after intervention had begun in only 8 sites, half in each arm.
  - Study population at high risk from COVID.
  - Facility dieticians busy with other duties.
  - High mortality in the 8 sites could complicate the primary outcome and reduce power.
  - Transfers from one facility to another could threaten internal validity.
  - Mortality and hospitalization could vary more across facilities than expected.
- Solutions
  - Intervention in other sites paused until September or later when the situation is more stable.
  - The number of facilities will be increased to guard against threats to power.

Guiding Good Choices for Health (GGC4H)

- Design: multi-center parallel group-randomized trial
- Groups: 24 pediatricians within each of 3 health care systems (g=36, m=50)
- Procedures: Adolescents age 12 recruited in two consecutive annual cohorts complete a baseline survey and annual follow-up surveys.
- Intervention: Pediatrician recommends Good Behavior Game intervention to parents in person or via mail. Staff offer parents group-based or self-guided intervention. Parents who cannot be reached or decline the group-based intervention receive the self-guided intervention. The group-based intervention sessions are limited to parents with the same pediatrician.
- Outcomes: annually via telephone survey with youth (incidence of any alcohol, tobacco, or marijuana use)
- Analysis: mixed model logistic regression with pediatrician as a nested random effect

GGC4H

- COVID struck as the first cohort of adolescents was being recruited.
  - Healthcare systems shifted most non-urgent care to virtual delivery.
  - Families have little appetite for non-essential, in-persons visits.
  - High-quality care can be delivered virtually, with high patient and provider satisfaction.
  - New rules allow billing for telehealth services.
- Solution
  - GGC is being delivered virtually; preserves 1 intervention, 2 modalities.
  - New delivery mode, but not a new intervention.
  - Group-format delivery via Zoom or similar technology.
  - Reduces some barriers to access: childcare, transportation.
  - Lack of computer or internet can be overcome via mobile phones.
  - Early data from health care systems suggest lower no show rates.
  - Design, data collection, data analysis unchanged.

Pragmatic Trial of Acupuncture for Chronic Low Back Pain in Older Adults (AcuOA)

- Design: 3-arm, multi-center individually randomized group-treatment (IRGT) trial with patients nested within acupuncturists or primary care providers
- Randomization: individual randomization after stratification by health care system, age, and gender (n=210)
- Intervention: 6 months acupuncture vs 3 months acupuncture vs usual care
- Outcomes: functional outcomes, pain intensity at baseline, 3, 6, and 12 months, with 6 months as the primary endpoint.
- Analysis: repeated measures GEE analysis with adjustment for baseline with standard errors obtained using the robust sandwich estimator.
  - The team expects each acupuncturist or PCP to have only a few patients, thereby limiting the impact of the nesting of patients within acupuncturists or PCPs.
  - They will perform an omnibus test of any difference among the 3 arms first, then move to pairwise comparisons if the omnibus test is significant.

Pragmatic Trial of Acupuncture for Chronic Low Back Pain in Older Adults (AcuOA)

- COVID struck during the UG3 phase of the trial, so it is still in planning.
  - Delayed pilot test of intervention as assessment protocols with associated extraction of EHR data.
  - Concerns about willingness of patient to visit acupuncturists during the pandemic.
- Solutions to date
  - Set power at 90%.
  - Recruit 813 to allow for 25% attrition.
  - Recruitment delayed until at least November.
  - The team is still working through details of the analytic plan
    - For example, considering whether to address missing data as part of the primary analysis to to address it in a secondary analysis.
    - Which covariates to include in addition to the baseline value.
    - Whether to model time as categorical or continuous.

General Approach to Addressing Disruptions

- Consider PICOTs
  - Consider each element in the PICOT framework and whether might be any COVID impact.
    - population, intervention, comparator, outcome, and time frame
- Consider CONSORT Guidelines
  - Consider each element in the methods and results section of the relevant CONSORT guidelines for the study design and whether there might be any COVID impact.
    - In addition to PICOTs, consider sample size, participant flow, recruitment, and additional secondary analyses
- Develop a plan for how the study will adapt in terms of additional measures, changes to the analytic plan, and trial reporting.
- Discuss the potential challenges and adaptations with the Collaboratory Biostatistics and Design Working Group.
- Implement the adaptations needed to preserve the integrity of the trial.

References / Resources

- *PICOT:*
  - Haynes et al. (2006) *Clinical epidemiology: how to do clinical practice research*.
  - Granger (2020) AACN Adv Crit Care

- *Consort Checklists:*
  - Schulz et al. (2010) Ann Intern Med (Randomized Controlled Trials)
  - Boutron et al. (2008) Ann Intern Med (Individually Randomized Group Treatment Trials)
  - Campbell et al. (2012) BMJ (Group-Randomized Trials)
  - Piaggio et al. (2012) JAMA (Noninferiority and Equivalence Trials)
  - Hemming et al. (2018) BMJ. (Stepped Wedge Group-Randomized Trials)