Cluster Randomized Trials and The Stepped Wedge Design

Megan Branda & Patrick Wilson

Cluster Design

- Unit of randomization other than the individual
  - Unit of analysis could be cluster or individual
  - Lack of independence among individuals within cluster

- When planning
  - Intervention
  - Eligibility Criteria (cluster and individual level)
  - Assessment of responses

Why cluster?

- Administrative convenience
- To obtain cooperation of investigators/administration
- Ethical considerations
- To enhance subject compliance
- To avoid treatment group contamination
- Intervention naturally applied at the cluster level

Randomization

- Completely randomized – no pre-stratification or matching of clusters
  - Best suited to large number of clusters
- Paired – two clusters in a stratum randomly assigned
  - Need consensus on what to pair on
  - Finding complete pairs for all clusters
  - Problem with estimate of intra – cluster correlation
- Stratified – Several clusters are randomly assigned
  - Number of clusters large enough in relation to number of factors stratifying by

Ethical considerations

- Is your trial 'minimal risk'?
  - 'the probability and magnitude of harm or discomfort anticipated in the research are not greater than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests.'
- Informed consent – from whom to you obtain…
  - Consent of subjects post randomization (subject not cluster)
  - Decision based upon intervention, data being collected, and who is considered a participant…
- Blinding participants
  - Most cases not feasible to blind clinicians/staff
  - Patients potentially

Sample size

- 'Intracluster Correlation Coefficient' (ICC) which is based on the relationship of the between to within-cluster variance
  - ICC calculated by a standard one-way analysis of variance among and within clusters
- Variance inflation factor (VIF): inflated by a factor $1 + (n - 1)\rho$
  - $\bar{n}$ is the average cluster size, and ρ is the estimated ICC (assuming the clusters are of a similar size)
- PASS calculates sample size for several different types of outcomes.

Unit of analysis?

- Cluster level:
  - Standard statistical approaches can be applied
  - Removes the problem of non-independence
  - Most appropriate when question of interest focuses on the unit of randomization
- Subject (patient) level:
  - Cluster adjusted techniques must be applied
    - Cluster adjusted chi-square, t-test; mixed effect linear regression modeling, GEE
    - Reduces to standard statistical approaches when cluster effect absent

Loss to follow-up

- Loss of cluster
  - If Paired randomization, then loss of one cluster means loss of its pair.
  - Some clusters back out once randomization is complete.
- Loss of patients
  - Relocation – change practice
  - High non-response
  - No shows to appointments
- Can you instead do a series of cross-sectional samples??

Considerations

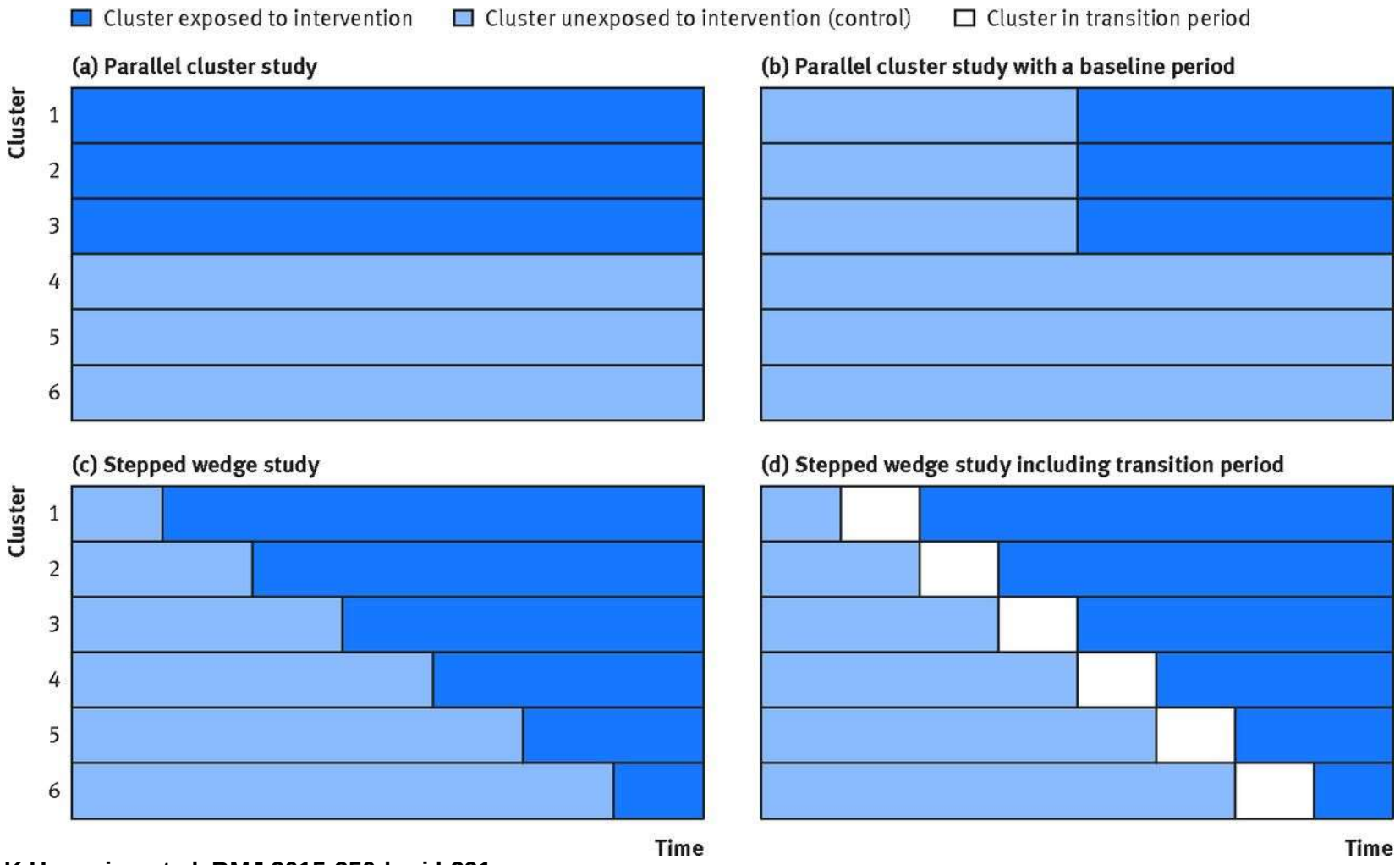
- Having consistent comparator can be problematic
  - Not all clinical practice is uniform
    - Amount of time with patient
    - Community or clinic resources
    - Ideological approach to care
  - Burnout of clinicians
- Potential for intervention variations between clusters as well as through implementation

Stepped Wedge Cluster Randomized Trials

Introduction to Design

- Clusters cross over from the control to treatment randomly
  - Not randomized *to* a treatment but rather randomized *when* they receive the treatment
- Cross over is staggered such that the intervention effect is partly, but not completely, confounded with time.
  - Requires additional modeling/assumptions due to this possible temporal confounding
- Although not inherent to the design generally done with cross-sectional samples

Schematic illustration of the conventional parallel cluster study (with variations) and the stepped wedge study.



K Hemming et al. BMJ 2015;350:bmj.h391

Statistical Model

- General form proposed by Hussey & Hughes

$$Y_{ijk} = \mu + \alpha_i + \beta_j + X_{ij}\theta + e_{ijk}$$

- Time can be modeled in several different ways
  - Each step having it's own effect
  - Parametric or semiparametric model
  - Conditional model (risk set)
- Can be solved in a Mixed model or GEE framework

Extended Designs

- Multiple Treatments
  - Concurrent, Supplementation, Replacement and Factorial
- Burn in, wash out periods
  - Add time points of no intervention
- Treatment heterogeneity
  - Add random effects
- Cohort designs
  - Additional layer of random effects

Power

- Powering is contingent on several factors
  - The number of clusters
  - Average size of cluster
  - The number of time points
  - Random cluster effect variance

- Two principle ways to power a study
  - Close form solutions
  - Simulation

Choices for Power

Close form/Design Effect

$$Var\left(\hat{\theta}\right) = \frac{6(T-1)(1-\rho)(\sigma_e^2+\tau^2)(1+(mT-1)\rho)}{mIT(T-2)\left(1+\left[\frac{m(T+1)}{2}-1\right]\rho\right)}$$

- Requires balance
- Can be extended to binary outcomes
- Also "design" effect formulas available

Simulation

- Select a average cluster size with a design matrix of treatments assigned to a fixed number of clusters and time points
- Provide an estimate of the main parameters
- Simulate a dataset of size n from the assumed model.
- Analyze the resulting dataset and record whether the intervention effect is detected as statistically significant.
- Repeat steps 3 & 4 for several iterations the proportion of statistically significant effects is your power

Why Wedge CRT vs. Standard CRT?

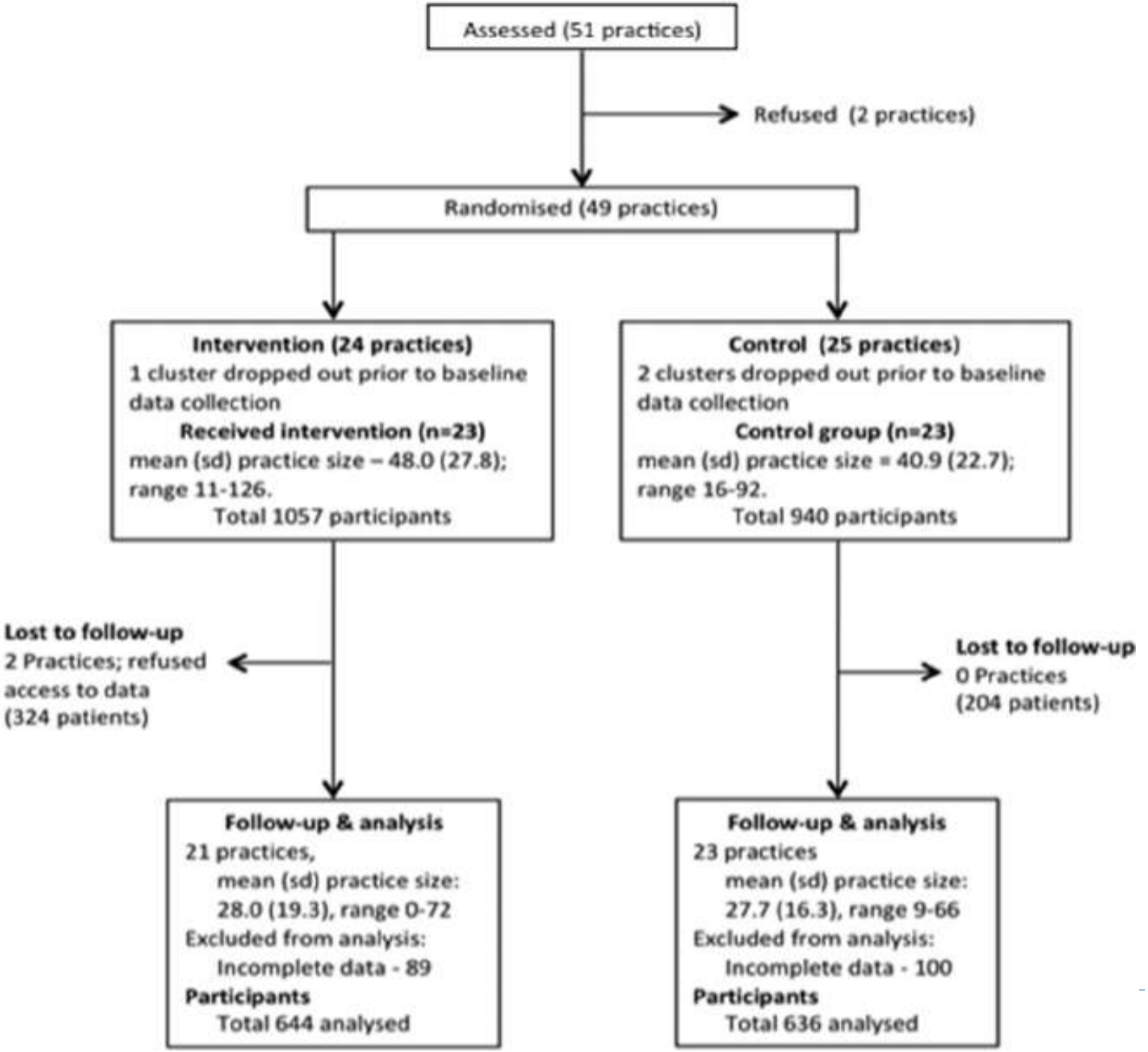
- Statistical considerations
  - Requires fewer clusters than a parallel design but more measurements per cluster
- Practical/Pragmatic issues
  - Population & measurements tend to be "real world"
  - Allows problems of feasibility, implementation to be addressed
- Political
  - Provides intervention to all clusters/communities
- Logistical
  - May reduce implementation costs through staggering training/start-up effort

References

- Current issues in the design and analysis of stepped wedge trials. James P. Hughes et. al.
- The Stepped Wedge Cluster Randomized Trial and Its Potential for Child Health Services Research: A Narrative review. Yasaman Fatemi and Robert Jacobson.
- Evaluation of multiple interventions using a stepped wedge design. Vivian Lyons et. al.
- Cluster Randomised Trials. Richard Hayes and Lawrence Moulton
- Design and Analysis of Cluster Randomization Trials in Health Research. Allan Donner and Neil Klar

How you report..

- Similar to CONSORT (extension available addressing cluster trials)
- Flow diagram expressed at cluster and individual level
- Demographics reported at individual and cluster level
- Outcome results accompanied by ICC


Assessed (51 practices)
Refused (2 practices)
Randomised (49 practices)
Intervention (24 practices)
1 cluster dropped out prior to baseline data collection
Received intervention (n=23)
mean (sd) practice size – 48.0 (27.8); range 11-126.
Total 1057 participants
Control (25 practices)
2 clusters dropped out prior to baseline data collection
Control group (n=23)
mean (sd) practice size = 40.9 (22.7); range 16-92.
Total 940 participants
Lost to follow-up
2 Practices; refused access to data
(324 patients)
Lost to follow-up
0 Practices
(204 patients)
Follow-up & analysis
21 practices,
mean (sd) practice size: 28.0 (19.3), range 0-72
Excluded from analysis:
Incomplete data - 89
Participants
Total 644 analysed
Follow-up & analysis
23 practices
mean (sd) practice size: 27.7 (16.3), range 9-66
Excluded from analysis:
Incomplete data - 100
Participants
Total 636 analysed