

Cluster Randomized Controlled Trials: Design and analysis issues

Developing an Integrated Strategy to Support
Pediatric and Perinatal Clinical Trials across Canada
West Brome, Quebec, May 27-29, 2011

Janusz Kaczorowski PhD
Director
Primary Care & Community Research, CFRI
Professor & Research Director
UBC Department of Family Practice

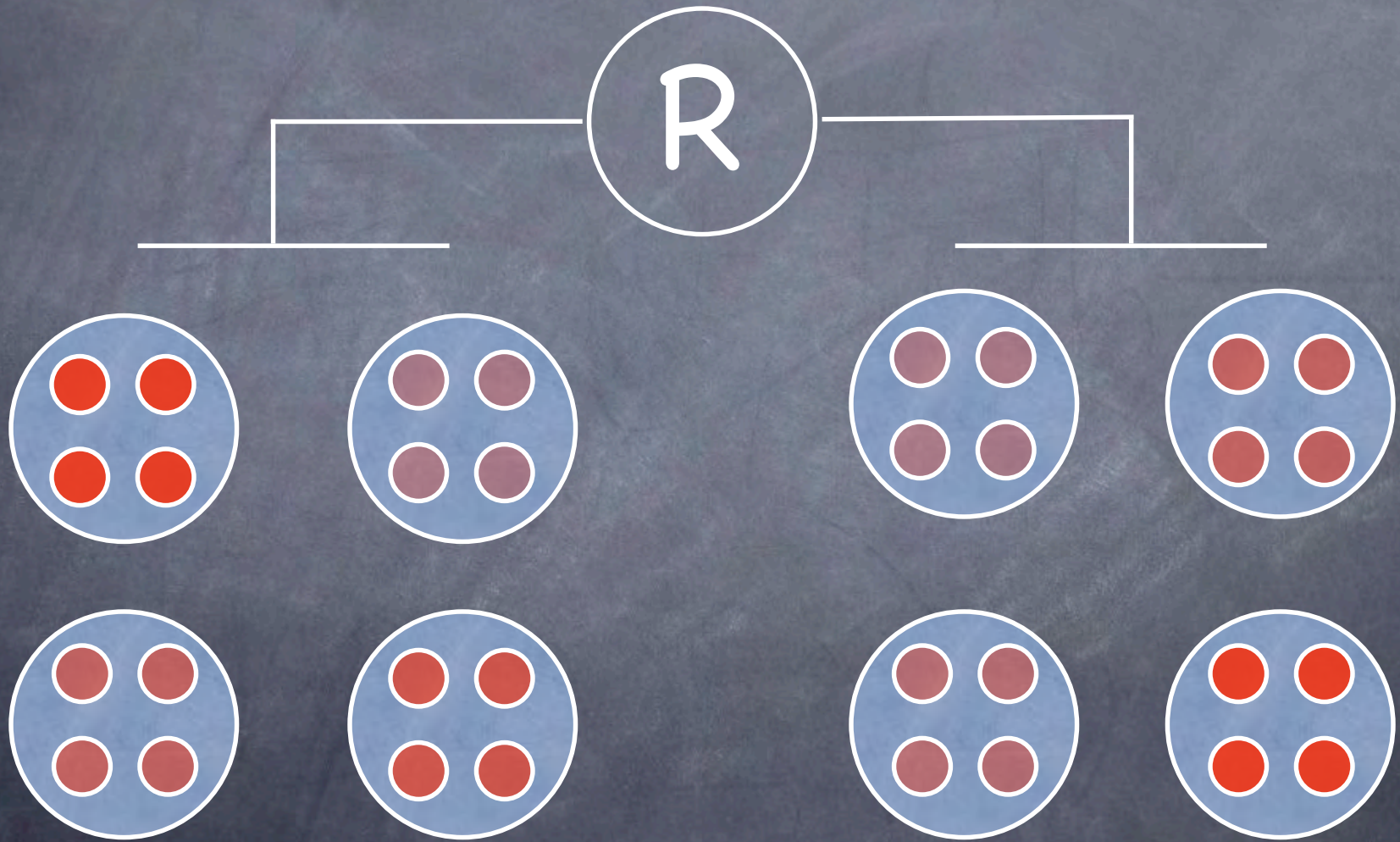
Outline

- What is a cluster RCT- why and when do we use it?
- Special considerations for sample size
- Special considerations for design
- Statistical analysis approaches
- Ethical considerations
- CHAP example

What is a cluster RCT?

- Cluster randomized trials are experiments in which (intact) social units or clusters rather than individuals are randomly allocated to intervention groups
 - Communities
 - Schools
 - Families
 - Hospitals/Clinics/Practices
- Increasingly used in health technology assessment studies

Clustering



Examples

Intervention	Social unit/cluster
Guideline implementation	Hospitals
Mass media campaign	Communities
Smoking prevention	Schools
Dietary intervention	Families

Challenges of cluster RCTs

- Outcome for each participant cannot be assumed to be independent of that for any other participant since those within the same cluster are more likely to have similar outcomes ('clustering effect')
- This lack of independence influences the RCT design and analysis
 - Standard approaches to sample size estimation and analysis are not appropriate

Impact on design & analysis

- Application of standard sample size formulas will lead to underpowered studies
 - To maintain the desired power, larger sample size is required
- Application of standard statistical methods will tend to bias p-values downward increasing spurious claims of statistical significance
 - To account/adjust for clustering special statistical methods are required

General Issues in Sample Size Estimation

- Issues common to sample size estimation that apply to any randomized trial:
 - Selection of the primary/secondary study outcomes
 - Determination of a minimally important effect of the intervention (effect size)
 - Specification of a statistical test or confidence interval method along with its directionality (one-sided/two-sided hypothesis)
- In addition, in cluster randomized trials:
 - Number and size of individual clusters
 - Prior assessment of intercluster correlation (ρ)

Intercluster correlation

- Because of the correlation of individual-level responses within clusters, there are two components of variation:
 - within cluster
 - between cluster
- The two components may be estimated either via the notion of intercluster correlation (ICC) or, a similar concept, the coefficient of variation between clusters (CV)

Measuring clustering effect

- Intercluster correlation coefficient (ICC) or rho (ρ): measure of the degree of similarity (correlation) of responses within a given cluster

$$\text{ICC or } \rho = \frac{s_b^2}{(s_b^2 + s_w^2)}$$

where s_b^2 = the variance between clusters
and s_w^2 = the variance within clusters
(from the F statistic from one way ANOVA)

Design Effect (DE)

- For sample size determination, "design effect" is defined as:
$$DE = 1 + \rho (m-1)$$
where ρ is the intercluster correlation and m is the cluster size
- DE is a measure of how much the sample size in each group have to be increased to achieve the same statistical power as would be obtained by individual level randomization
- When $\rho = 0$, $DE = 1$ and the responses within clusters are independent

Example

- Sample size to detect 20% difference (40% vs 60%) using standard formula with a continuity correction (power = 80%; two-sided alpha = 0.05), $n = 214$ (107 per group)
- Assuming that these 214 subjects are recruited from 20 clinics (~ 10 subjects per clinic) and ICC = 0.1
 - $DE = 1 + \rho (m-1)$; $1 + 0.1 (10-1) = 1.9$
- This means that, in order to maintain our original power, the sample size needs to be increased by 90% (new $n = 407$)

Balance between 'k' (# of clusters) and 'm' (size of clusters)

- As 'k' increases (# of clusters), power increases
- Increasing 'm' (size of clusters) has little impact on power
 - One can never get rid of clustering effect by increasing 'm'
- If clusters are very small, clustering effect has much less impact on power even if ICC is substantial
 - i.e. If we were to recruit only 1 patient from 214 different practices, our original power would hold regardless of how large/small ICC was

Sample size considerations

- Availability/feasibility: balance between number of clusters and participants within each cluster
- Baseline outcome proportion or mean (SD)
- Effect size desired
- Intercluster correlation coefficient (ICC) ie.
 - $\text{Variance within Ss} / \text{Total variance}$

Design issues:

Randomization Options

- Randomization is done at the cluster level (all subjects within a given cluster are allocated to the same intervention arm)
- Randomization choices are critical as investigators are usually confronted with a relatively small number of clusters
- Completely randomized (unrestricted)
- Restricted
 - Stratified
 - Matched pair

Randomization: Stratification

- Fear that simple randomization might not be effective given small number of clusters
- However, number of potential strata is usually quite limited
- Cluster size itself should be considered as a stratifying factor
- Debate as to whether strata should be subsequently modeled as fixed or random factors

Other design issues: Post-randomization recruitment bias

- After clusters are randomized, individuals may need to be recruited
 - Potential bias in how participants are selected if recruiter knows allocation status
- Potential solution- recruit clusters and individuals within these clusters, then randomize clusters

Other design issues: Concealment & blinding

- Lack of concealment may be a source of selection bias:
 - Solution: develop objective measures of eligibility and ensure that the allocation is carried out by someone who is independent of the trial
- Participant blindness (clinics, physicians, patients) to intervention assignment (one of the pillars of validity) is difficult or impossible to maintain in cluster RCTs
 - Ensure that outcome assessors are blinded
 - Comply with CONSORT for cluster RCTs

Cluster Analysis Options

- Cluster level analysis
 - T-test (or weighted t-test) at cluster level of analysis
- Patient level analysis
 - Adjusted statistics (chi-square, t, F, z values)
 - Random effects meta-analysis approach (only for paired cluster design)
 - Mixed linear models, hierarchical linear models, generalized estimating equations (GEE)

T-test: Cluster level analysis

- Calculate a summary measure for each cluster, such as a cluster mean or proportion
- Because each cluster then provides only one data point, the data can be considered to be independent, allowing standard statistical tests such as t-test to be used
- If the size of the clusters varies widely, use a weighted t-test, using cluster sizes as the weights
- Generally, not statistically efficient

Adjusting for design effect

- Adjusting test statistics for design effect
 - chi-square or F-test statistics divided by the design effect
 - t-test or z-test statistics divided by the square root of the design effect
 - Brier's chi-square
 - Rosner and Milton's chi-square
 - Donner and Donald's chi-square
 - Rao and Scott's chi-square

Generalized estimating equations (GEE)

- Extension of logistic regression that adjusts for effects of clustering
- Does not require parametric assumptions
- Tests odds ratio of intervention effect
- Can adjust for individual and cluster level covariates
- Requires approximately 40 clusters to be valid
- Must specify the correlation structure but fairly robust if specification not correct

Multivariate techniques

- Multivariate modelling techniques allow adjustment for clustering and for both cluster level and patient level covariates
- The a priori model-fitting analysis strategy should identify:
 - covariates which are to be considered for inclusion
 - order in which confounding variables are to be considered for inclusion in the model (with the intervention variable fitted last)

Ethical Considerations

- All the same principles apply as in individually randomized RCTs
- Who gives consent when individual participants are not recruited?
- Individual consent may be theoretically possible but not feasible
- If individual consent not possible, there must be a “cluster representation mechanism”
 - Clusters have the same rights as an individual participant would have- including withdrawal of cluster from the study

Ethical consideration

- Two types of cluster RCTs:
 - Interventions that are received (or not) by a whole cluster together – fluoridation of the water supply in the intervention communities
 - Interventions where individuals can decide individually, without reference to others, to receive it or not

C-CHAP example: trial objective

- To evaluate the effectiveness of CHAP in reducing stroke/CVD morbidity at the community level:
- **Primary outcome measure:** hospital admissions for acute myocardial infarction, congestive heart failure, and stroke (composite end-point) among ALL residents aged ≥ 65 years
- **Design:** community cluster RCT
- **Data sources:** routinely-collected, population-based administrative health data

Inclusion/exclusion

- Inclusion criteria:

- Community size: 10,000 – 60,000
- Number of family physicians: 5+
- Number of pharmacies: 2+
- Total community-dwelling population: 65+

- Exclusion criteria:

- Immediately adjacent to metro area (e.g. Dundas)
- Rural /dispersed (e.g. townships & native reserves)
- Participated in CHAP demonstration project (e.g. Grimsby & Brockville)

Study Flowchart



CHAP intervention

- Community-wide promotion of CHAP sessions (letters from FPs, referrals and local media campaigns)
- Trained peer volunteers help participants to measure and record BP with accurate, automated device (BPTru™) and fill out standardized CVD and stroke risk profile
- BP and risk factor information captured via fax-to-database technology and shared with family physicians, pharmacists and participants
- Participants receive education materials and links to local/provincial/national resources targeted to specific modifiable risk factors
- Community health nurse and pharmacist available to assess participants with high BP

CHAP implementation

- All 20 randomly selected communities successfully launched CHAP
- 214/341 physicians 'actively participated'
- 24,196 personalized invitation letters mailed
- 129/145 pharmacies participated
- 577 volunteers recruited & trained
- 1,265 sessions held
- 27,358 assessments (15,889 unique participants)
- ~25% of older adults in CHAP communities attended at least one CHAP pharmacy session



Baseline characteristics

Measure	Control (n=19)	CHAP (n=20)
No. of residents aged 65+	3 829.89 ± 2 176.44	3 393.70 ± 1 831.59
Age (in years)	74.79 ± 0.43	74.82 ± 0.62
% Male	42.65 ± 1.19	42.92 ± 2.16
Rurality Index	28.96 ± 13.60	31.63 ± 14.09
% Low income status	16.95 ± 8.55	18.57 ± 11.33
No. of prescription drugs	7.25 ± 0.49	6.98 ± 0.54
No. of Comorbidity Groups	7.31 ± 0.30	7.17 ± 0.50
Charlson Comorbidity Index	0.57 ± 0.09	0.58 ± 0.11
% with diabetes	22.16 ± 2.34	21.20 ± 2.79
% with history of CHF	12.19 ± 1.91	12.45 ± 2.34
Death rate per 100	3.45 ± 0.40	3.55 ± 0.57

Hospital admission rates per 1,000

Outcome	Before CHAP n=67 874	Before Control n=72 768	After CHAP n=69 942	After Control n=75 499	Rate Ratio (95% CI)
Composite	30.15	29.36	27.90	30.13	0.91 (0.86-0.97) p<0.01
AMI	10.24	10.26	9.54	10.81	0.87 (0.79-0.97) p<0.01
CHF	11.19	11.11	10.51	12.22	0.90 (0.81-0.99) p=0.03
Stroke	8.71	7.99	7.86	7.10	0.99 (0.88-1.12) p=0.89

Interpreting RR = 0.91

- Extrapolating these results to the population 65+ in Ontario, UK and USA would result in approximately 5 000, 30 000, and 120 000 fewer annual CVD hospital admissions, respectively
- On par with the benefits of population-wide reductions in dietary salt (2g/day reduction), tobacco use (elimination of 40% of use of or exposure to tobacco), or obesity (5% BMI reduction in obese individuals) on annual number of CVD events

Key points to remember

- Rationale for adopting a cluster design including randomization scheme
- Incorporating the effects of clustering into the sample size calculations
- Incorporating the effects of clustering into the analysis
- Developing flow of both clusters and individuals through the trial, from assignment to analysis
- Addressing special ethical issues

Additional resources

- Cluster randomised trials: Methodological and ethical considerations:
<http://www.mrc.ac.uk/Utilities/Documentrecord/index.htm?d=MRC002406>
- CONSORT: extension to cluster trials:
<http://www.consort-statement.org/extensions/designs/>
- Pitfalls of and Controversies in Cluster Randomization Trials (Donner A, Klar N: Am J Public Health. 2004;94:416-422)
- Analysis of cluster randomized trials in primary care: a practical approach. (Campbella M et al: Family Practice Vol. 17, No. 2, 192-196)