MRC | Biostatistics Unit

A cluster-randomised cross-over trial

Design of Experiments in Healthcare
Isaac Newton Institute, Cambridge
15th August 2011

Ian White
MRC Biostatistics Unit, Cambridge, UK

Plan

1. The PIP trial
2. Why cluster-randomise?
3. Why cross-over?
4. Efficiency
5. Analyses
6. How widely useful is the cluster-crossover design?

1. The Parent-Baby Interaction Programme

- Pre-term birth (<32 weeks) places a child at high risk of neurological and psychological impairment
  - possibly because the psychological stress associated with preterm birth adversely affects mother-infant interaction
- Hence the "Parent-Baby Interaction Programme" (PBIP), a nurse-led intervention to enhance the parent-infant relationship & help parents care appropriately for their preterm babies
- Hope for positive effects on infant cognitive outcomes

The Pre-term Infant Parenting (PIP) trial

- The PIP trial evaluated the Parent-Baby Interaction Programme in 6 neonatal centres in England
  - 233 babies recruited from 210 mothers
  - each centre admitted ~ 3 pre-term babies per month
  - mean length of stay for each baby was ~ 2 months
  - recruitment was spread over a year
- Intervention was delivered by specially trained research nurses
- Comparator was usual care
- Outcomes included
  - parental stress at 3 months
  - child's mental and physical development at 24 months

2. PIP: why cluster-randomise?

Two main reasons to cluster-randomise:

- To avoid contamination:
  - with ~ 6 pre-term babies in each centre at any one time, an individually-randomised trial would allow transfer of knowledge about the intervention from intervention to control parents
- For practical reasons: e.g.
  - intervention delivered in fewer centres
  - better compliance

Problems with cluster-randomising

- Possible recruitment bias
  - recruitment must take place over time (as the babies are born) so must follow centre randomisation
  - risk of differential consent between the two arms
  - we addressed this
    - » *by trying to maximise consent rates (achieved 233/307=76%)*
    - » *by monitoring consent rates to check they were similar across arms (80% intervention, 72% control)*
- With only 6 centres, power is likely to be low
  - addressed by the cross-over part of the design (see next)
- Analysis must allow for clustering

3. PIP: why cross-over?

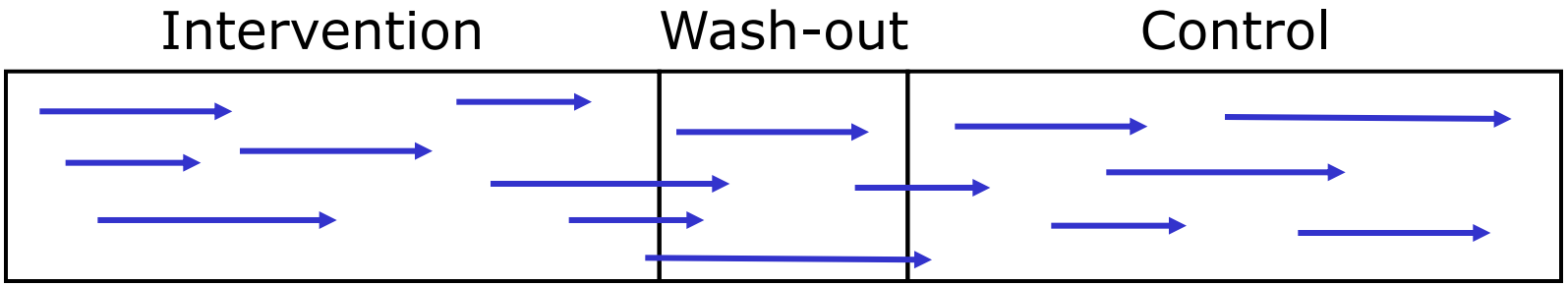
- Cross-over trials are used in individually randomised designs to reduce the impact of between-individual variation (hence improving precision)
- Since clusters are likely to vary in their average outcomes, the same argument applies
- We allocated each centre to one intervention for 6 months, then the other intervention for 6 months
  - order was randomised
- Allocation applied to all babies born in the centre within the 6-month period

Problems with cross-over design

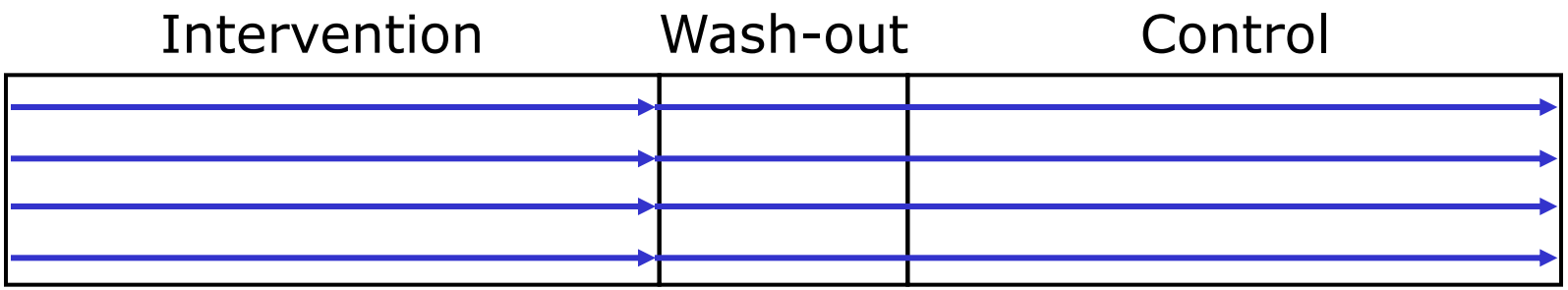
- Possible carry-over: treatment in one period affects outcome in later periods
  - in individually-randomised drug trials: drug still in system, or physiological or psychological effects remain
  - in cluster-randomised trial like PIP: giving the intervention in period 1 could affect later outcomes through presence of the same individuals (parents & staff?)
  - avoid carry-over by having an adequate wash-out period
  - PIP: a 3-month wash-out separated the two periods
- Analysis must allow for pairing

Two cluster-randomised cross-over designs

Relatively short stays, different patients in each period:

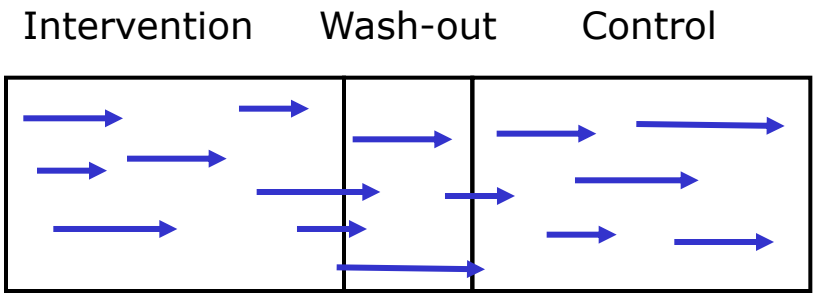


Long stays, same patients measured in each period:



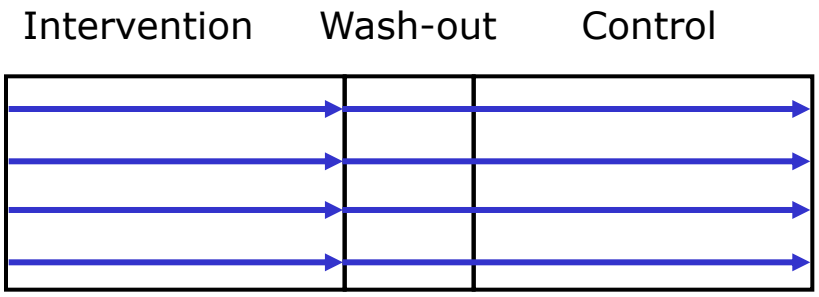
Carry-over

Short stays



- Need to avoid carry-over at institutional level
- e.g. by providing intervention through dedicated staff who aren't present in control period

Long stays



- Need to avoid carry-over at institutional and individual levels
- Only suitable for chronic diseases?

4. Efficiency

- A model for a CRXO trial:
  - let x_{ij} = treatment allocated to cluster i in period j
  - let y_{ijk} = outcome for individual k in period j in cluster i
- Model

$$y_{ijk} = \alpha_j + \beta x_{ij} + u_i + v_{ij} + e_{ijk}$$

$$u_i \sim N(0, \sigma_u^2)$$

$$v_{ij} \sim N(0, \sigma_v^2)$$

$$e_{ijk} \sim N(0, \sigma_e^2)$$

- Here β is the intervention effect of primary interest

Efficiencies of 3 designs

3 designs with Kn individuals per arm:

- IRT: individually randomised trial with Kn individuals per arm, 1 period
- CRT: cluster-randomised trial with K clusters of size n per arm, 1 period
- CRXO: cluster-randomised cross-over trial with K/2 clusters of size n per arm, 2 periods
- We can compute efficiencies with equal cluster sizes
  - for the estimated intervention effect is simply the difference of the means

Efficiencies of 3 designs

Model $y_{ijk} = \alpha_j + \beta x_{ij} + u_i + v_{ij} + e_{ijk}$

IRT: $\text{var}(\hat{\beta}) = \frac{1}{K}\left(\frac{\sigma_u^2}{n} + \frac{\sigma_v^2}{n} + \frac{\sigma_e^2}{n}\right)$

CRT: $\text{var}(\hat{\beta}) = \frac{1}{K}\left(\sigma_u^2 + \sigma_v^2 + \frac{\sigma_e^2}{n}\right)$

CRXO: $\text{var}(\hat{\beta}) = \frac{1}{K}\left(\sigma_v^2 + \frac{\sigma_e^2}{n}\right)$

CRT always inferior to IRT

CRXO superior to IRT

if $\sigma_v^2 < \sigma_u^2/(n-1)$

CRXO superior to CRT

if $\sigma_u^2 > 0$

NB: fewer degrees of freedom of CRXO trial also matters.

Two intra-cluster correlations (ICCs)

ICC between periods (same cluster):

$$\rho_{\text{period}} = \sigma_v^2 / \left(\sigma_v^2 + \sigma_e^2 \right)$$

ICC between clusters:

$$\rho_{\text{cluster}} = \left(\sigma_u^2 + \sigma_v^2 \right) / \left(\sigma_u^2 + \sigma_v^2 + \sigma_e^2 \right)$$

Note $\rho_{\text{period}} \leq \rho_{\text{cluster}}$

5. CRXO trial: analysis questions

- Should we use cluster-level analyses?
  - based on cluster means
  - easier to understand
- or individual-level analyses?
  - based on hierarchical models
- With equal cluster sizes (i.e. each cluster recruits n individuals in each period), all methods give the same point estimates
  - & all are unbiased even with unequal cluster sizes
- Main interest is in
  1. Standard errors → coverage
  2. Efficiency with unequal cluster sizes

Individual-level analyses considered

- Model $y_{ijk} = \alpha_j + \beta x_{ij} + u_i + v_{ij} + e_{ijk}$
- v_{ij} , e_{ijk} always random
- M1: random cluster effects u_i
- M2: fixed cluster effects u_i
- M2(-): M1 and M2 constrain all random effect variances to be non-negative, but M2(-) has σ^2_v unconstrained
- All fit by REML

Cluster-level analyses considered

- Adapted from analysis of a cross-over trial
- Define d_i = "cross-over difference" for cluster i:
  mean in intervention period – mean in control period
- Define $w_i = x_{i2} - x_{i1}$ = +1 if int→cont, -1 if cont→int
- Intervention effect is β in model $d_i = \beta + \gamma w_i + e_i$

- Unweighted: fit model by OLS
  - estimate: ½{ave d in int→cont + ave d in cont→int}
- Weighted 1: fit model with weight = $n_{i1}\, n_{i2} / (n_{i1}+n_{i2})$
- Weighted 2: weight that also allows for ICCs
- Weights only matter with unequal cluster sizes

Simulation study: coverages with 6 clusters of equal sizes

20 individuals per cluster-period

ICC period	0	0	0	0.05	0.05	0.1
ICC cluster	0	0.05	0.1	0.05	0.1	0.1
M1	99.5%	99.4%	99.7%	97.4%	97.5%	96.4%
M2	99.6%	99.4%	99.7%	98.0%	97.5%	96.9%
M2(–)	94.6%	94.2%	94.2%	94.8%	95.1%	95.1%
Cluster-level (all methods)	94.6%	94.2%	94.2%	94.8%	95.1%	95.1%

Forcing variance components ≥ 0 → over-coverage.

Cluster-level methods implicitly allow negative variance components.

Simulation study: empirical standard errors with 6 clusters of different sizes

Cluster-period sizes = Poisson(μ), μ = 15 or 25

ICC period	0	0	0	0.05	0.05	0.1
ICC cluster	0	0.05	0.1	0.05	0.1	0.1
M1	1.32	1.31	1.34	1.85	1.91	2.35
M2	1.33	1.32	1.34	1.85	1.91	2.35
M2(−)	1.34	1.33	1.35	1.86	1.92	2.36
Cluster						
unweighted	**1.39**	**1.36**	**1.39**	**1.88**	**1.93**	**2.36**
n-weighted	1.32	1.31	1.34	1.86	1.92	2.37
using ICC	1.33	1.32	1.34	1.85	1.91	2.35

Unweighted cluster method has *small* loss of efficiency.

Simulation study: conclusions

- We analysed the trial using cluster-level methods with $n_{i1}\ n_{i2} / (n_{i1}+n_{i2})$ weights
- We adjusted for individual-level baseline variables z by applying the cluster-level analysis to the residuals from an OLS regression of outcome on z
- But we assessed interactions (of intervention with individual-level covariates) using a hierarchical model

An awkward analysis point

- In a preliminary analysis, we found a negative estimated variance component, and hence the standard error allowing for clustering was lower than not allowing for clustering
- Felt that we would want to report the larger standard error in this case
- In other words we prefer M2 to M2(-)
- Correct coverage is not the most important thing?

- This situation can arise in all CRTs

6. Summary: advantages & disadvantages of a CRXO design

Advantages

- Efficiency
- Convenience

Disadvantages

- Recruitment bias is always a threat?
  - can be avoided in 1-period CRTs by listing & consenting all individuals before randomisation, but this is unlikely in a 2-period CRT
  - can arise if eligibility is under staff control (PIP: if staff could influence baby's date of birth – unlikely)
  - can arise through differential consent
- Analysis is somewhat complicated

When is the cluster-crossover design useful?

- Other examples of its use:
  - compare 2 policies for ordering chest X-rays for ventilated intensive care patients (Lancet 2009; 374: 1687-93)
    - » *is carry-over a risk?*
  - evaluate real-time audio-visual feedback about cardio-pulmonary resuscitation performed outside hospital (BMJ 2011; 342: d512)
    - » *clusters alternated between feedback-on and feedback-off over 2-5 periods*
- Seems ideal for evaluating policy-type interventions which can be switched on and off, provided recruitment bias can be avoided

Acknowledgements & references

- PIP trial: Sam Johnson, Cris Glazebrook, Chrissie Israel, Andrew Whitelaw, Neil Marlow
  - Glazebrook C, Marlow N, Israel C, Croudace T, Johnson S, White IR, Whitelaw A. Randomised trial of a parenting intervention during neonatal intensive care. *Archives of disease in childhood* 2007; **92**: 438–43.
  - Johnson S, Whitelaw A, Glazebrook C, Israel C, Turner R, White IR, Croudace T, Davenport F, Marlow N. Randomised trial of a parenting intervention for very preterm infants: Outcome at 2 years. *Journal of Pediatrics* 2009; **155**: 488–94.
- Methods project: Rebecca Turner, Tim Croudace
  - Turner RM, White IR, Croudace T. Analysis of cluster randomised cross-over trial data: a comparison of methods. *Statistics in Medicine* 2007; **26**: 274–289.