Choosing appropriate analysis methods for cluster randomised cross-over trials with a binary outcome

Katy E Morgan

London School of Hygiene and Tropical Medicine

28th March 2017

What is a cluster randomised cross-over trial?

Cluster randomised trials:

- randomise groups rather than individuals.
- have to account for dependency in sample size calculation and analysis.

Cross-over trials:

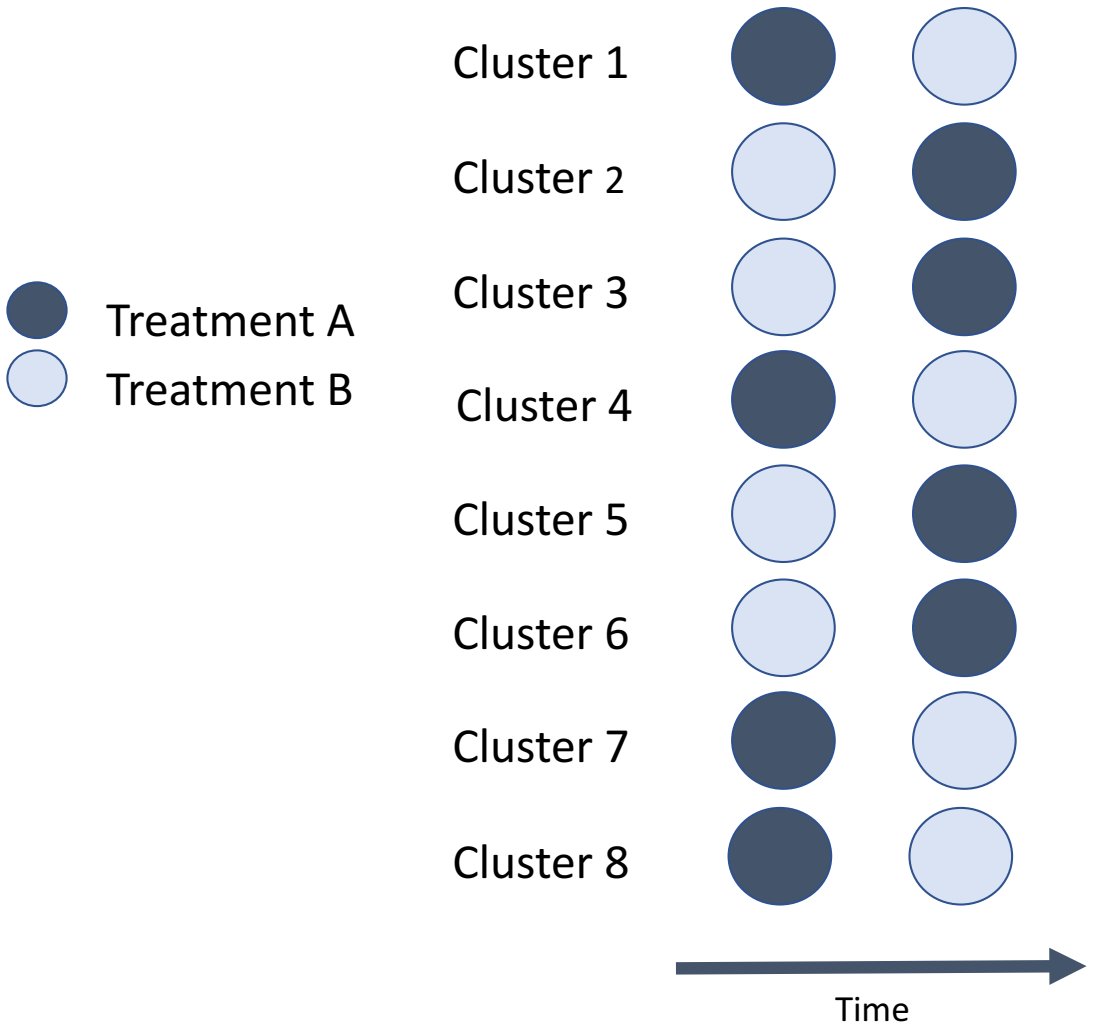
- randomised to A/B or B/A.
- subjects act as their own control.
- have clustering of measurements within subjects.

What is a cluster randomised cross-over trial?

Cluster randomised cross-over (CRXO) trials:

- **Groups** of individuals receive **multiple** treatments. Order is randomised.
- Can have same individuals in all periods, different individuals, or a mixture.

What is a cluster randomised cross-over trial?



Example:

- 2 treatments
- 2 treatment periods
- 8 clusters

Why might they be used?

- May have to use cluster randomisation (e.g. intervention acts at group-level).
- Lose power by doing this.

Why might they be used?

- May have to use cluster randomisation (e.g. intervention acts at group-level).
- Lose power by doing this.

- In a CRXO, have within cluster comparison which may **reduce** required size of a cluster randomised trial.
- But do have to consider chance of treatment carry-over.

Why is the analysis important?

Two sources of clustering:

- correlations in **clusters**
- correlations in **period within cluster**

$\rightarrow$ Leads to a more complex analysis.

Not handling correlations appropriately could lead to **incorrect or misleading results**.

Systematic review

- Arnup *et al.* (2016) systematic review of 139 analyses from 91 CRXO trials.
- Found that only 10% (14/139) analyses used potentially appropriate methods that account for both cluster and cross-over design elements.

Systematic review

- Arnup *et al.* (2016) systematic review of 139 analyses from 91 CRXO trials.
- Found that only 10% (14/139) analyses used potentially appropriate methods that account for both cluster and cross-over design elements.

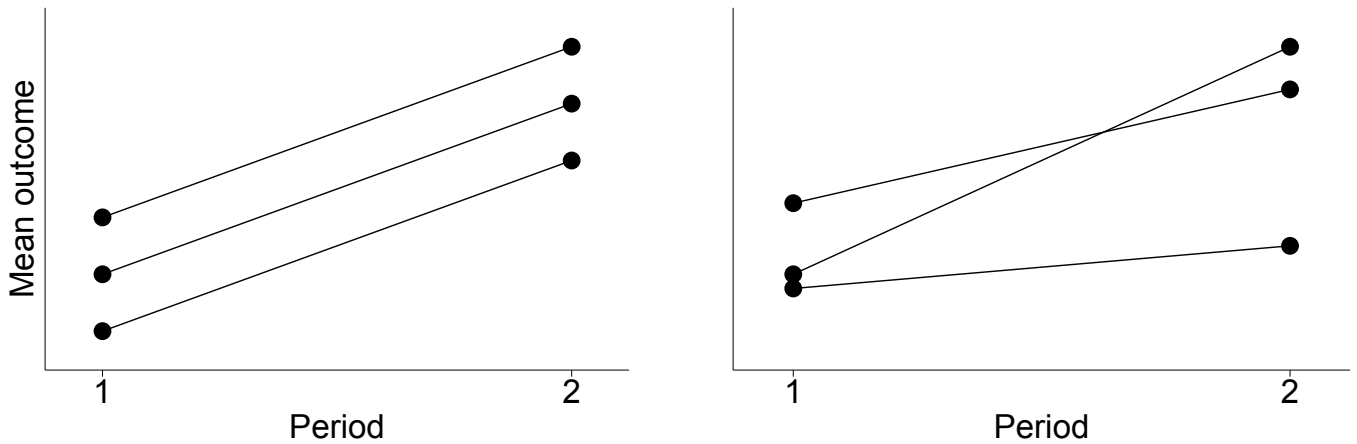
Morgan, Forbes, Keogh, Jairath and Kahan, *Stat. Med.* (2017):

- Simulation study for binary outcomes from CRXO trial
- Two-period, two-treatment trial design
- Different patients in each period

Simulations: methods of analysis

Logistic hierarchical models:

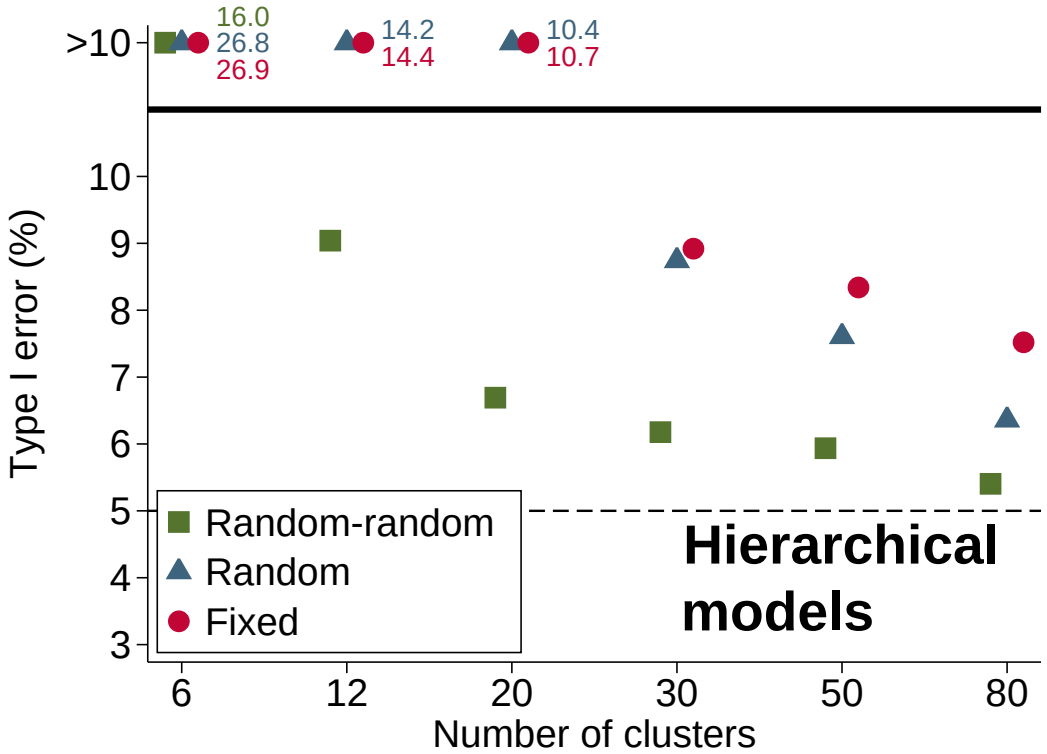
- **Fixed** cluster effects,
- **Random** cluster effects,
- **Fixed** cluster effects – **random** period within cluster effects,
- **Random** cluster – **random** period within cluster effects.



Cluster-level summary method: linear regression on proportion of events in each period within cluster.
+ various GEEs and weighted cluster-level summary regressions.

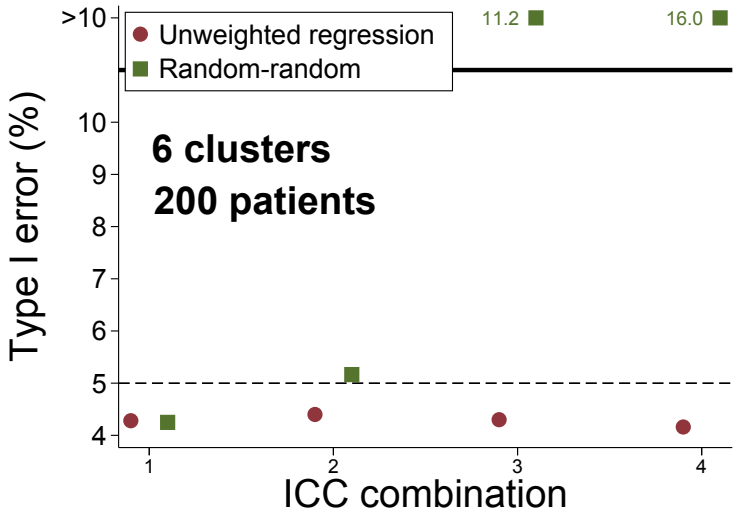
Simulations

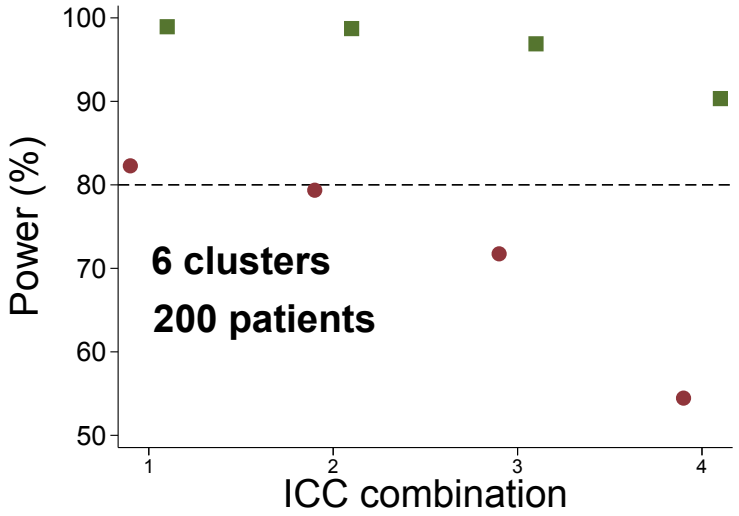
- Generated data from a random-random hierarchical model.
- 5000 data sets per scenario.



- Fixed-random has convergence problems.
- Hierarchical models — Type I inflated to over 10% for scenarios with few clusters and extra within-period correlation.
- Cluster effects only perform worse than random–random.

Factorial simulations





- Different combinations of within-cluster and within-period correlations.
- Random-random hierarchical model has inflated Type I errors for scenarios with extra within-period correlation and small numbers of clusters.
- Unweighted cluster-level summary regression has good Type I error, but loses power when correlations are high.

Conclusions

- CRXO design can be useful in some settings to increase power compared to a cluster-randomised trial.
- But have more complex analysis.
- Simulations for binary outcome, two-period CRXO trial with different participants:
  - Ignoring extra within-period correlation can lead to inflated Type I errors: should account for this in analysis method.
  - Need to have a large number of clusters to use a random-random hierarchical model.
  - Have to consider potential loss of power if using a cluster-level summary method with a small number of clusters — consider this at design stage.

References

Morgan, Forbes, Keogh, Jairath, Kahan, *Stat Med.* (2017) 36: 318–333

Arnup, Forbes, Kahan, Morgan, McKenzie, *J Clin Epi.* (2016) 74: 40–50

Forbes, Akram, Pilcher, Cooper, Bellomo, *Clin Trials* (2015) 12: 34–44

Turner, White, Croudace, *Stat Med.* (2007) 26: 274-289

Rietbergen, Moerbeek, *JEBS* (2011) 36: 472–490

Systematic review

- Median number of clusters 9 (IQR 4-21).
- 58 trials (69% of those with number of periods available) had only 2 periods.
- 27 trials (30%) included same individuals in all periods.
- Only 9% (12/139) analyses performed at cluster-level.
- Out of 127 individual-level analyses, only 4 used potentially appropriate methods. 54 did not account for clustering or cross-over elements. No analyses used random effect for cluster-period.

Simulation parameters

Initial simulation — increasing number of clusters

- 15% event rate in control arm, first period
- Fixed period effect OR of 0.85
- Cluster ICC 0.062, extra cluster-period ICC 0.023 ($\sigma_c^2 = 0.137, \sigma_p^2 = 0.081$)
- Number of clusters 6–80, number of patients per cluster-period 200–8

Simulation parameters

Initial simulation — increasing extra period within cluster correlation

- 15% event rate in control arm, first period
- Fixed period effect OR of 0.85
- Cluster ICC 0.062
  - extra cluster-period ICC 0.001 ($\sigma_c^2 = 0.214, \sigma_p^2 = 0.003$)
  - extra cluster-period ICC 0.005 ($\sigma_c^2 = 0.200, \sigma_p^2 = 0.017$)
  - extra cluster-period ICC 0.01 ($\sigma_c^2 = 0.182, \sigma_p^2 = 0.035$)
  - extra cluster-period ICC 0.05 ($\sigma_c^2 = 0.042, \sigma_p^2 = 0.176$)
- Number of clusters 6 or 30, number of patients per cluster-period 200 or 22 respectively

Simulation parameters — factorial simulations

- 15% or 45% event rate in control arm, first period
- Fixed period effect OR of 0.85 (15% event rate) or 0.92 (45% event rate)
- No treatment effect or OR 0.5 (for 15% event rate) or OR 0.75 (for 45% event rate)
- ICC combinations (cluster ICC, extra cluster-period ICC)
  - 0.023, 0 ($\sigma_c^2 = 0.077, \sigma_p^2 = 0$)
  - 0.062, 0 ($\sigma_c^2 = 0.217, \sigma_p^2 = 0$)
  - 0.023, 0.01 ($\sigma_c^2 = 0.044, \sigma_p^2 = 0.034$)
  - 0.062, 0.023 ($\sigma_c^2 = 0.137, \sigma_p^2 = 0.081$)
- For 15% event rate: 6 clusters, number of patients per cluster-period 200 or 330; or 30 clusters, number of patients per cluster-period 22 or 31
- For 45% event rate: 6 clusters, number of patients per cluster-period 400 or 600; or 30 clusters, number of patients per cluster-period 55 or 75
  Note: corresponds to 80%, 90% power for ICC combination 2

Simulation parameters — further simulations

Random–random hierarchical model only

- 15% event rate in control arm, first period
- Fixed period effect OR of 0.85
- No treatment effect or OR 0.5
- ICC combinations (cluster ICC, extra cluster-period ICC)
  - 0.023, 0 ($\sigma_c^2 = 0.077, \sigma_p^2 = 0$)
  - 0.062, 0 ($\sigma_c^2 = 0.217, \sigma_p^2 = 0$)
  - 0.023, 0.01 ($\sigma_c^2 = 0.044, \sigma_p^2 = 0.034$)
  - 0.062, 0.023 ($\sigma_c^2 = 0.137, \sigma_p^2 = 0.081$)
- Number of clusters 6–100, number of patients per cluster-period 200–6

Back-up slide: Simulation results

- If no extra period within cluster correlation, then size-weighted cluster-level summary method works well.
- BUT Type I error is inflated for high values of extra period within cluster ICC.
- Same for ICC weighted regression, plus sometimes give negative weights.
- None of the GEE models considered in initial simulations had appropriate Type I errors.