Stepping up and crossing over: Designing efficient cluster randomised trials

Jessica Kasza: jessica.kasza@monash.edu

MATRIX online seminar series, November 2021

The Cluster Cluster at Monash

Andrew Forbes

Kelsey Grantham

Ehsan Rezaei

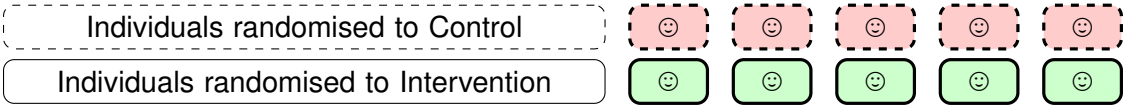
Stephane Heritier

Jess Kasza

Rhys Bowden

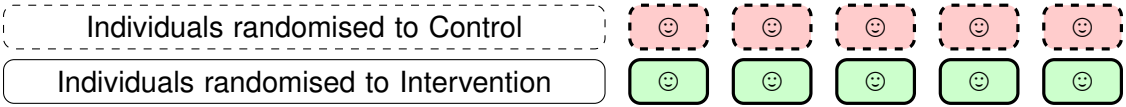
Lizzie Korevaar

Individually randomised trials



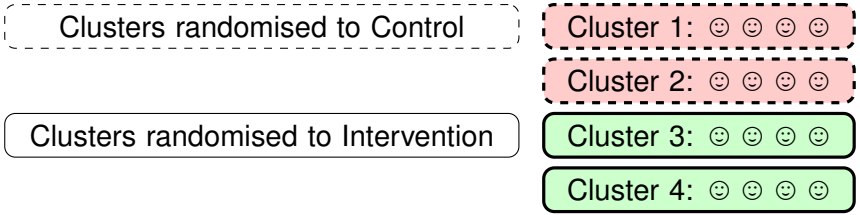
- Randomisation implies that we expect control and intervention groups to be similar
- Individual randomisation is the “gold standard” for assessing the effect of interventions
  - But it isn’t always possible!

Individually randomised trials



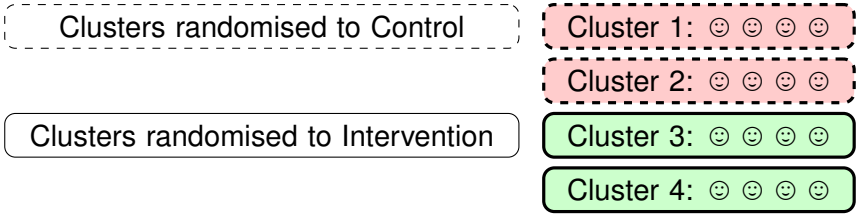
- Randomisation implies that we expect control and intervention groups to be similar
- Individual randomisation is the “gold standard” for assessing the effect of interventions
  - But it isn’t always possible!
  - Sometimes interventions are applied at a “group” level.

The standard cluster randomised trial



- Intervention at group level may mean that we can't randomise individuals!
  - Instead we need to randomise *clusters (groups)* of participants to treatments.
- Clusters could be hospitals, intensive care units, schools, neighbourhoods...
  - The **intracluster correlation (ICC)** (ρ) describes the similarity of outcomes from participants in the same cluster.

The standard cluster randomised trial



- Intervention at group level may mean that we can't randomise individuals!
  - Instead we need to randomise *clusters (groups)* of participants to treatments.
- Clusters could be hospitals, intensive care units, schools, neighbourhoods...
  - The **intracluster correlation (ICC)** (ρ) describes the similarity of outcomes from participants in the same cluster.

The greater this similarity, the more participants we need to get the same amount of information about the intervention effect!

The large impact of small ICCs

Even a small ICC can lead to a large number of required clusters.

- "Design effect"': for clusters of size m

$$1 + \rho(m - 1)$$

- E.g. $m = 100$, $\rho = 0.01$, design effect = 1.99.
- Diminishing returns when including extra participants in each cluster: will need to add more clusters instead.

The large impact of small ICCs

Even a small ICC can lead to a large number of required clusters.

- "Design effect"': for clusters of size m

$$1 + \rho(m - 1)$$

- E.g. $m = 100$, $\rho = 0.01$, design effect = 1.99.
- Diminishing returns when including extra participants in each cluster: will need to add more clusters instead.

Cluster randomised trials not always feasible:

- What if more clusters are required than are available?

Extending cluster randomised trials over time

Longitudinal cluster randomised trials can help!

- Clusters are followed over time and measured repeatedly.
  - Participants within clusters may be measured only once (in one period), or contribute measurements in more than one study period.
- Clusters may switch between control and intervention conditions.

Extending cluster randomised trials over time

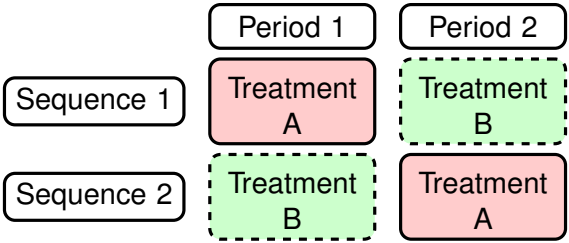
Longitudinal cluster randomised trials can help!

- Clusters are followed over time and measured repeatedly.
  - Participants within clusters may be measured only once (in one period), or contribute measurements in more than one study period.
- Clusters may switch between control and intervention conditions.

Switching between control and intervention conditions is really useful!

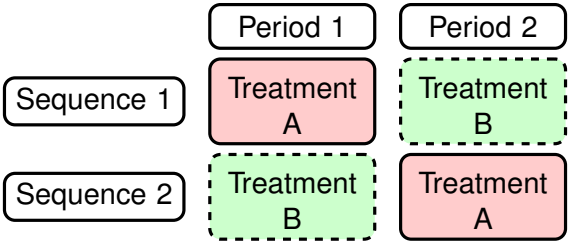
- Allows comparisons not only between clusters, but also within clusters
- ... which can lead to increases in power!

The cluster randomised cross over trial: CRXO



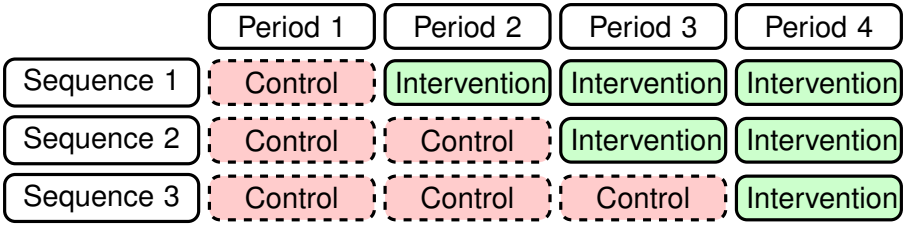
- Each cluster switches: treatment effect is now estimated using **BOTH** within-cluster and between-cluster comparisons.
- The most efficient design you can have!

The cluster randomised cross over trial: CRXO



- Each cluster switches: treatment effect is now estimated using **BOTH** within-cluster and between-cluster comparisons.
- The most efficient design you can have!
- But there may be practical reasons why the CRXO can't be used...
  - Switching back and forth is not always possible.
  - Might be rolling out a change across clusters.

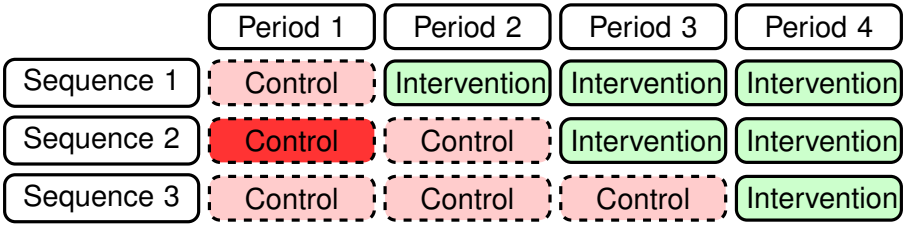
The stepped wedge cluster randomised trial design



Stepped wedge designs are wonderful!

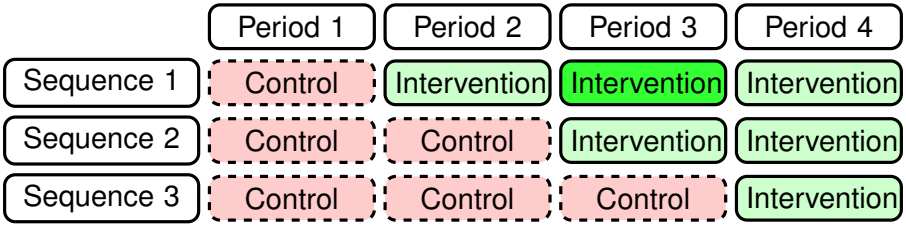
- Useful in situations where we can't switch back and forth between treatments.
- All clusters know they will receive the intervention (eventually...);
- Useful when interventions cannot be undone or will be rolled out anyway.

Does each *cell* contribute the same amount of information?



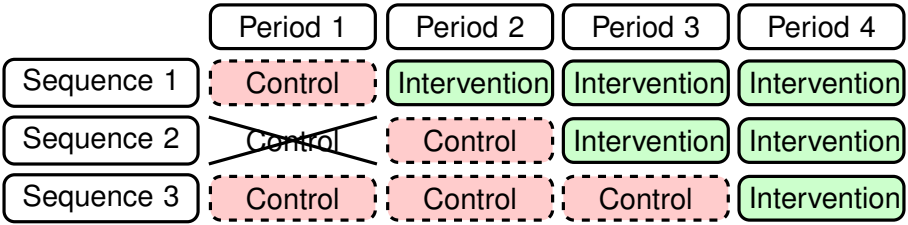
- Which participants contribute the most information about the treatment effect?
- Which cells can be omitted with the smallest acceptable decrease in power (or precision)?

Does each *cell* contribute the same amount of information?



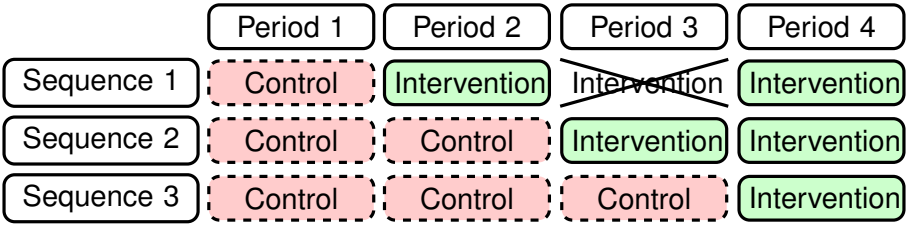
- Which participants contribute the most information about the treatment effect?
- Which cells can be omitted with the smallest acceptable decrease in power (or precision)?

Does each *cell* contribute the same amount of information?



- Which participants contribute the most information about the treatment effect?
- Which cells can be omitted with the smallest acceptable decrease in power (or precision)?

Does each *cell* contribute the same amount of information?



- Which participants contribute the most information about the treatment effect?
- Which cells can be omitted with the smallest acceptable decrease in power (or precision)?

Models for continuous outcomes

A simple model for continuous outcomes:

$$\text{Outcome} = \text{Period effect} + \textbf{Treatment effect} + \text{random effects} + \text{error}$$

Models for continuous outcomes

A simple model for continuous outcomes:

$$\text{Outcome} = \text{Period effect} + \textbf{Treatment effect} + \text{random effects} + \text{error}$$

For participant $i = 1, \ldots, m$, in period $t = 1, \ldots, T$, in cluster $k = 1, \ldots, K$:

$$Y_{kti} = \beta_t + X_{kt}\theta + \gamma_{kt} + \epsilon_{kti}, \quad \epsilon_{kti} \sim N(0, \sigma_\epsilon^2)$$

$$\gamma_k = (\gamma_{k1}, \ldots, \gamma_{kT}) \sim N_T(0, \mathbf{V}_\gamma)$$

X_{kt} is the treatment indicator for cluster k in period t.

Models for continuous outcomes

A simple model for continuous outcomes:

$$\text{Outcome} = \text{Period effect} + \textbf{Treatment effect} + \text{random effects} + \text{error}$$

For participant $i = 1, \ldots, m$, in period $t = 1, \ldots, T$, in cluster $k = 1, \ldots, K$:

$$Y_{kti} = \beta_t + X_{kt}\theta + \gamma_{kt} + \epsilon_{kti}, \quad \epsilon_{kti} \sim N(0, \sigma^2_\epsilon)$$

$$\gamma_k = (\gamma_{k1}, \ldots, \gamma_{kT}) \sim N_T(0, \mathbf{V}_\gamma)$$

X_{kt} is the treatment indicator for cluster k in period t.

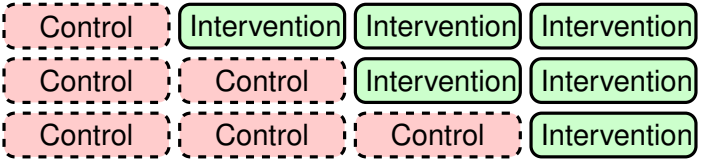
θ is the treatment effect: this is what we want to estimate.

- $\hat{\theta}$ the generalised least squares estimator of the treatment effect θ.
- $var(\hat{\theta})$ of interest: used in sample size calculations.

How much does $var(\hat{\theta})$ increase if observations from a given cell are omitted?

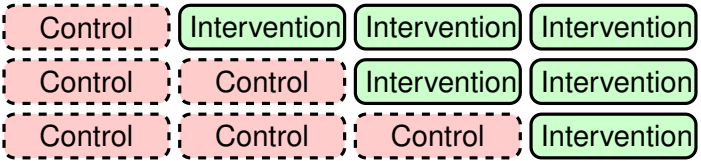
Information content of each cell

Calculate $var(\hat{\theta})$ given the complete design:

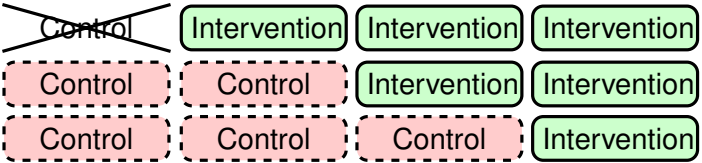


Information content of each cell

Calculate $var(\hat{\theta})$ given the complete design:

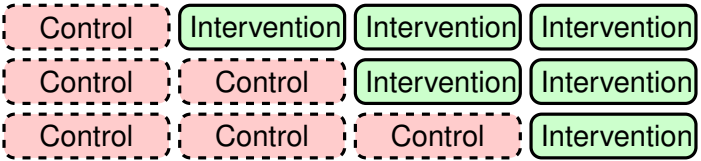


Calculate $var(\hat{\theta})_{[kt]}$ from the incomplete design, omitting period t of cluster k:

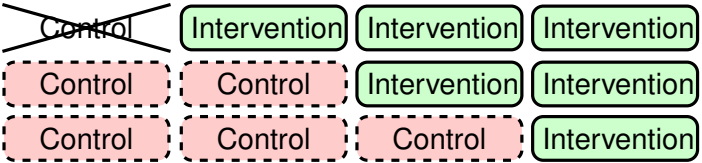


Information content of each cell

Calculate $var(\hat{\theta})$ given the complete design:



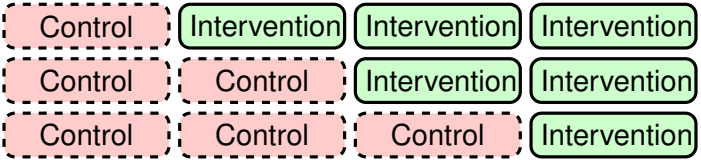
Calculate $var(\hat{\theta})_{[kt]}$ from the incomplete design, omitting period t of cluster k:



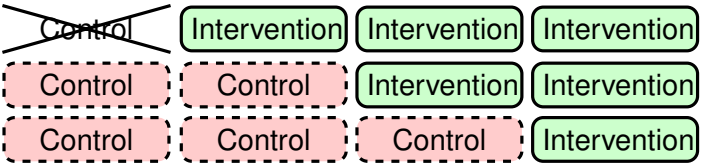
Information content of cell (k, t)**:** $IC(k,t) = var(\hat{\theta})_{[kt]}/var(\hat{\theta})$

Information content of each cell

Calculate $var(\hat{\theta})$ given the complete design:



Calculate $var(\hat{\theta})_{[kt]}$ from the incomplete design, omitting period t of cluster k:



Information content of cell (k,t)**:** $IC(k,t) = var(\hat{\theta})_{[kt]}/var(\hat{\theta})$

$IC(k,t) = 1$ implies no information loss; $IC(k,t) > 1$ implies loss of information.

Information content of cells: outcome models

$var(\hat{\theta})$ depends on the **within-cluster correlation structure**.

$$Y_{kti} = \beta_t + X_{kt}\theta + \gamma_{kt} + \epsilon_{kti}, \quad \epsilon_{kti} \sim N(0, \sigma_\epsilon^2)$$
$$\gamma_k = (\gamma_{k1}, \ldots, \gamma_{kT}) \sim N_T(0, \mathbf{V}_\gamma)$$

$\mathbf{V}_\gamma$ is the covariance matrix for the random effects

- We will consider the three most frequently-used structures, and what these say about correlations between subjects in the same cluster...
  - in the same or in different periods.

Statistical models for Y_{kti}: subject i; period t; cluster k.

- Model 1: constant correlation/exchangeable

$$Y_{kti} = \beta_t + \theta X_{kt} + \alpha_k + \epsilon_{kti}, \quad \alpha_k \sim N(0, \sigma_\alpha^2), \quad \epsilon \sim N(0, \sigma_\epsilon^2)$$

Statistical models for Y_{kti}: subject i; period t; cluster k.

- Model 1: constant correlation/exchangeable

$$Y_{kti} = \beta_t + \theta X_{kt} + \alpha_k + \epsilon_{kti}, \quad \alpha_k \sim N(0, \sigma^2_\alpha), \quad \epsilon \sim N(0, \sigma^2_\epsilon)$$

- Model 2: block exchangeable.

$$Y_{kti} = \beta_t + \theta X_{kt} + \alpha_k + \gamma_{kt} + \epsilon_{kti}, \quad \alpha_k \sim N(0, \sigma^2_\alpha), \quad \gamma_{kt} \sim N(0, \sigma^2_\gamma), \quad \epsilon \sim N(0, \sigma^2_\epsilon)$$

Statistical models for Y_{kti}: subject i; period t; cluster k.

- Model 1: constant correlation/exchangeable

$$Y_{kti} = \beta_t + \theta X_{kt} + \alpha_k + \epsilon_{kti}, \quad \alpha_k \sim N(0, \sigma^2_\alpha), \quad \epsilon \sim N(0, \sigma^2_\epsilon)$$

- Model 2: block exchangeable.

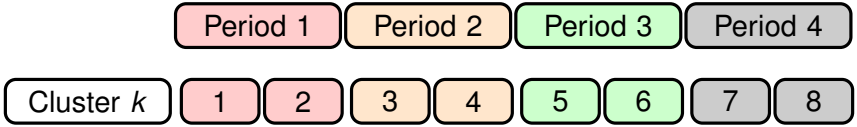
$$Y_{kti} = \beta_t + \theta X_{kt} + \alpha_k + \gamma_{kt} + \epsilon_{kti}, \quad \alpha_k \sim N(0, \sigma^2_\alpha), \quad \gamma_{kt} \sim N(0, \sigma^2_\gamma), \quad \epsilon \sim N(0, \sigma^2_\epsilon)$$

- Model 3: discrete-time decay.

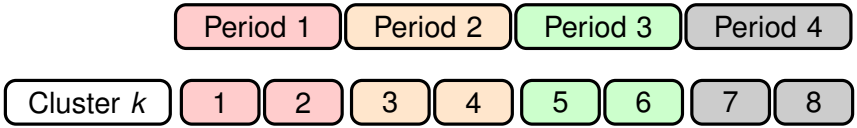
$$Y_{kti} = \beta_t + \theta X_{kt} + \gamma_{kt} + \epsilon_{kti}, \quad \epsilon \sim N(0, \sigma^2_\epsilon)$$

$$\gamma_k = (\gamma_{k1}, \ldots, \gamma_{kT}) \sim N(\mathbf{0}, \sigma^2_\gamma \Sigma), \quad \Sigma[t, s] = r^{|t-s|}$$

Four periods, 2 participants in each cluster in each period



Four periods, 2 participants in each cluster in each period



- **Within-period ICC**: correlation between a pair of participants in the same period;
- **Between-period ICC**: correlation between a pair of participants in different periods.
- Consider the correlation between participant 1 and each other participant in cluster k.

Within-cluster correlation structure: Model 1

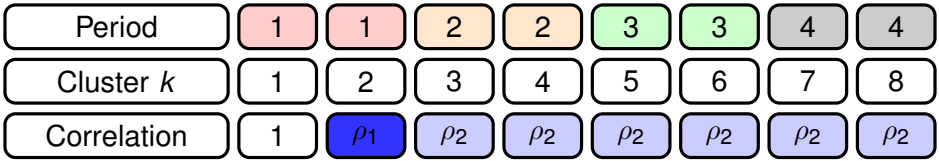
Period	1	1	2	2	3	3	4	4
Cluster k	1	2	3	4	5	6	7	8
Correlation	1	ρ	ρ	ρ	ρ	ρ	ρ	ρ

Model 1: Constant correlation. Within-period ICC = between-period ICC

- Correlation between the outcomes of any pair of participants is identical.
  - Correlation between participant 1 and 2 = correlation between 1 and 8.

NOTE: Shading indicates the degree of correlation between participant 1 and other participants. The correlation of participant 1 with themselves is 1.

Within-cluster correlation structure: Model 2



Model 2: "Block-exchangeable". Separate within- and between-period ICCs

- Participants in the *same* treatment period have more highly correlated outcomes than participants in different treatment periods.

NOTE: Shading indicates the degree of correlation between participant 1 and other participants. The correlation of participant 1 with themselves is 1.

Within-cluster correlation structure: Model 3

Period	1	1	2	2	3	3	4	4
Cluster k	1	2	3	4	5	6	7	8
Correlation	1	ρ	ρr	ρr	ρr^2	ρr^2	ρr^3	ρr^3

Model 3: Discrete-time decay

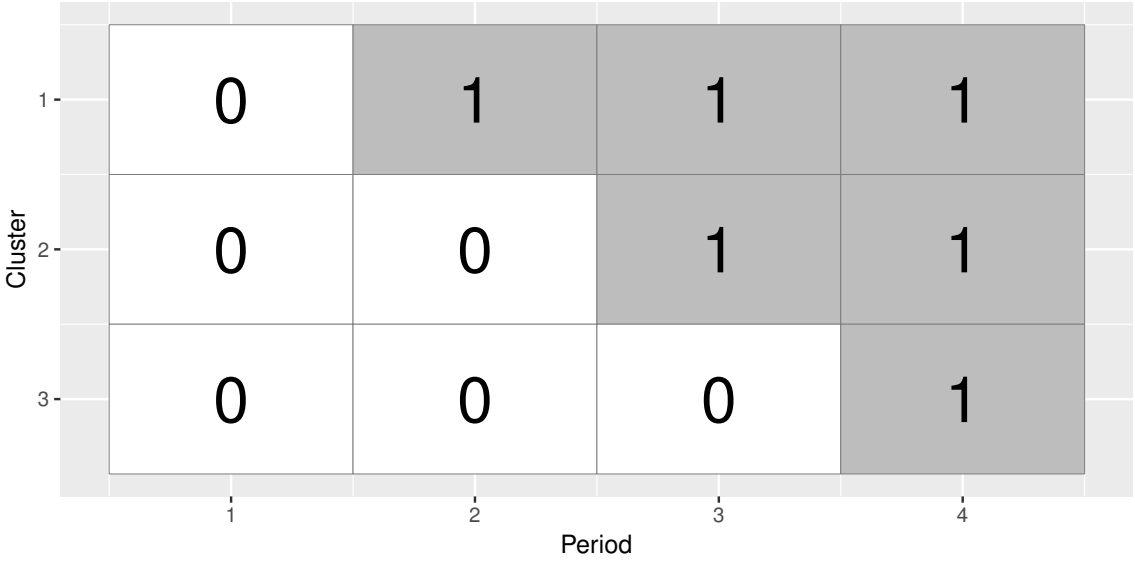
- The correlation between participant outcomes depends on their periods.
  - The correlation between a pair of participants decreases the further their measurement periods are apart in time.

NOTE: Shading indicates the degree of correlation between participant 1 and other participants. The correlation of participant 1 with themselves is 1.

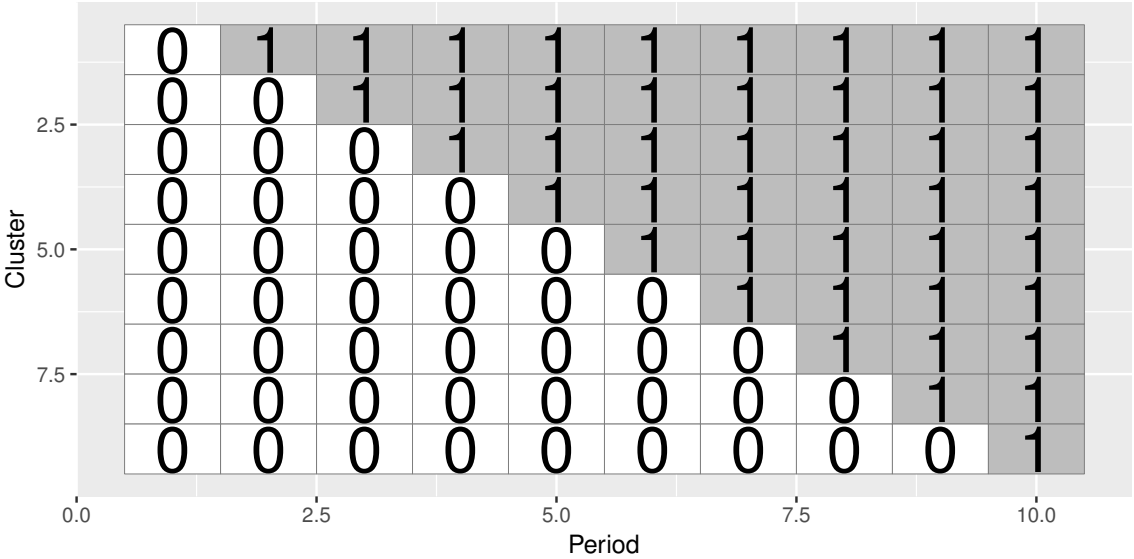
Investigating the information content of stepped wedge cells

- Consider stepped wedge designs with 4 and 10 periods.
  - One cluster per sequence.
  - 100 participants in each cluster-period.
- Model 1 with $\rho = 0.05$
- Model 3 with $\rho = 0.05$, $r = 0.95$

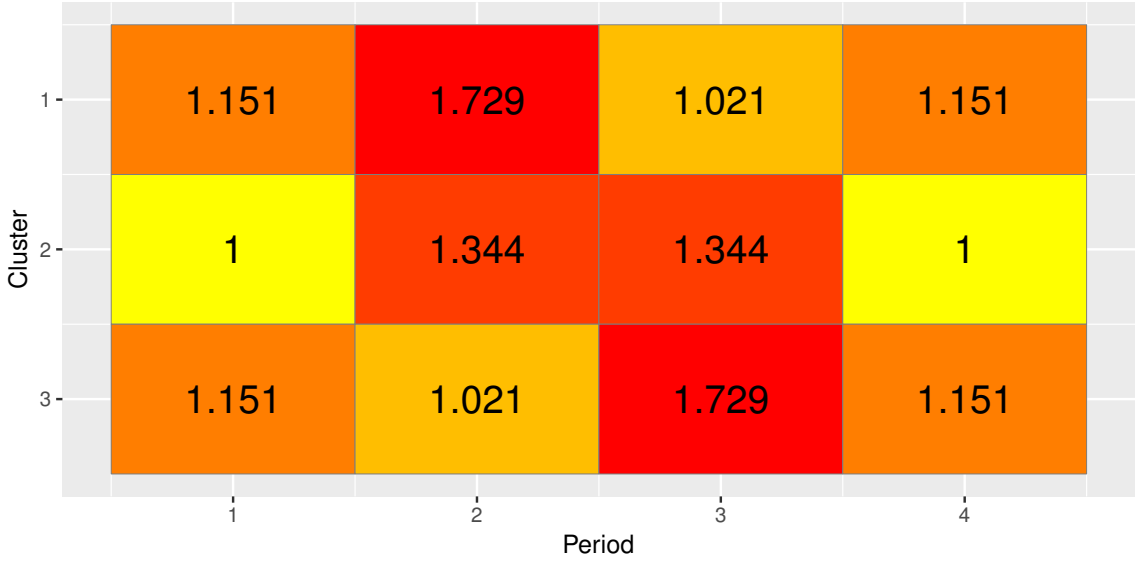
Design matrix: $T = 4$



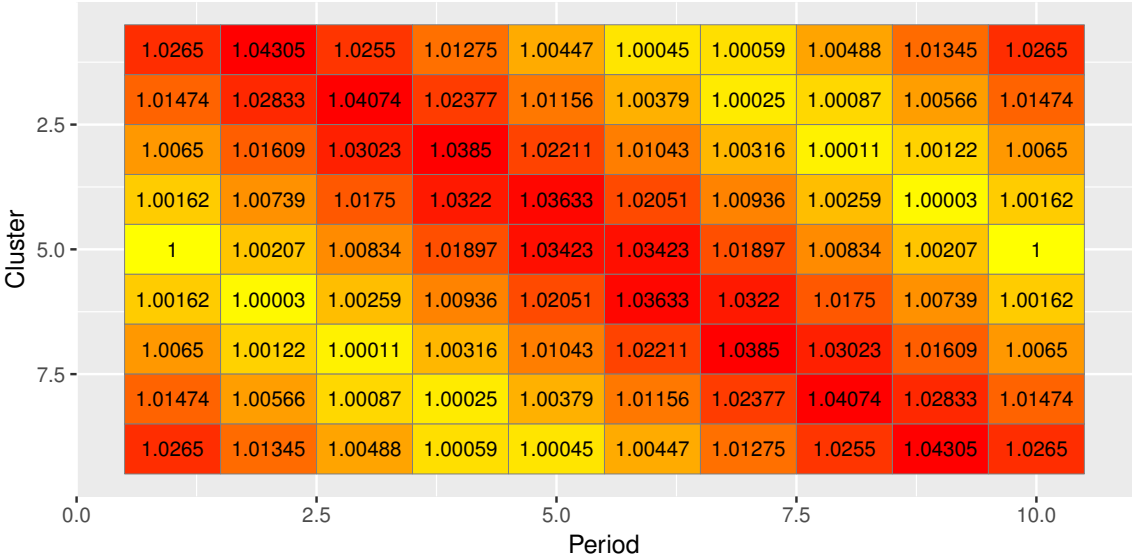
Design matrix: $T = 10$



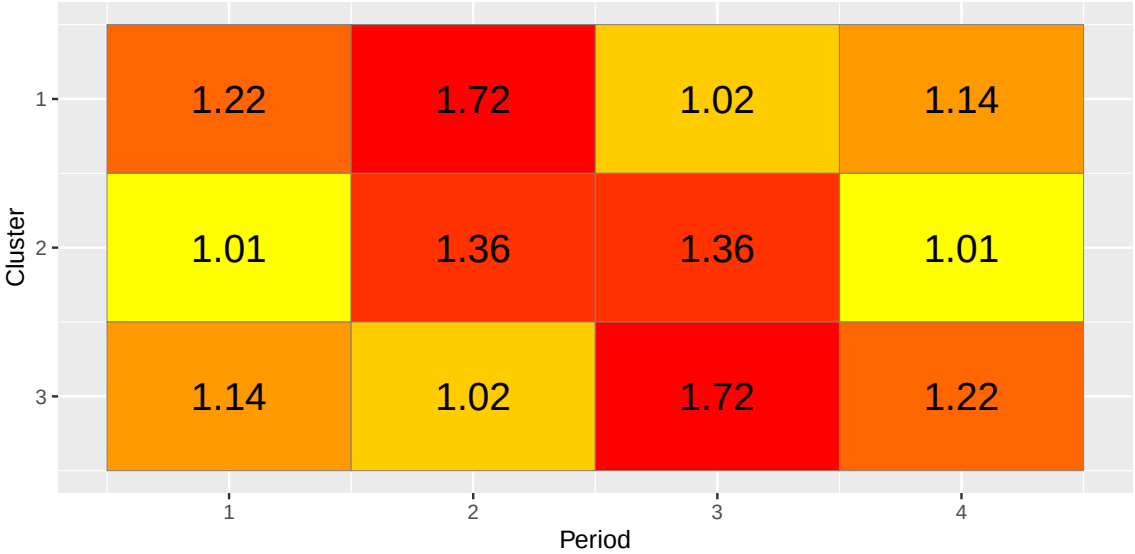
Information content of cells, Model 1: $\rho = 0.05$, $m = 100$



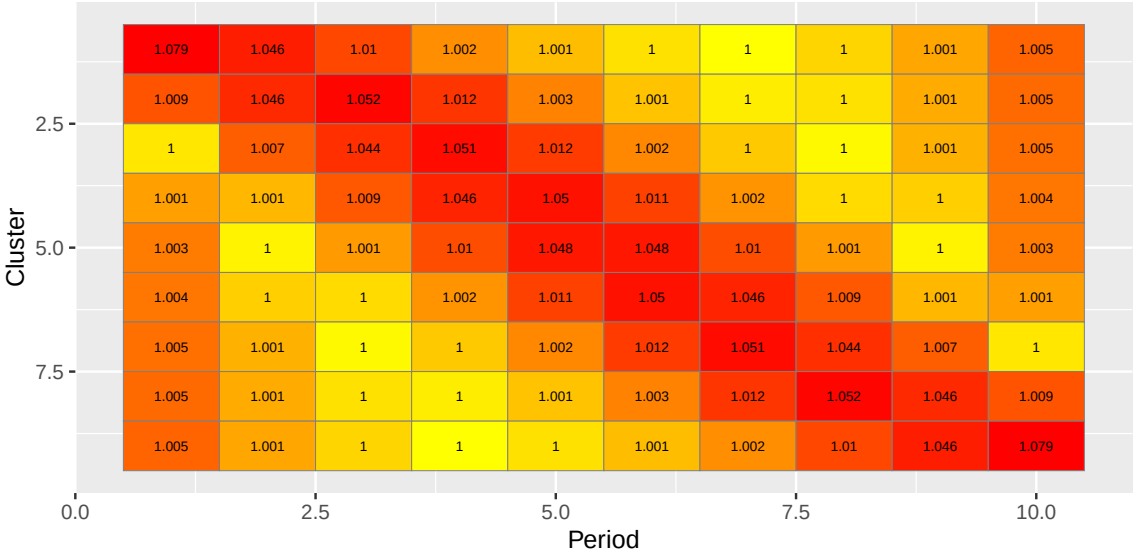
Information content of cells, Model 1: $\rho = 0.05$, $m = 100$



Information content of cells, Model 3: $\rho = 0.05$, r = 0.95, $m = 100$



Information content of cells, Model 3: $\rho = 0.05$, r = 0.95, $m = 100$



Information content of cells: theoretical results

For Models 1, 2, and 3 we get the following property:

Centrosymmetry: $IC(k,t) = IC(K+1-k, T+1-t)$

Further, for Models 1 and 2:

Information-free cells: $IC\left(\frac{K+1}{2}, 1\right) = IC\left(\frac{K+1}{2}, T\right) = 1$

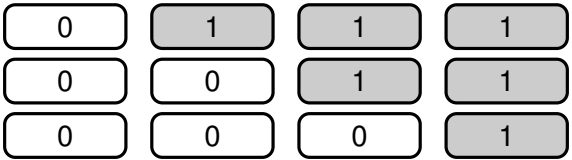
Information content of cells: theoretical results

For Models 1, 2, and 3 we get the following property:

$$\textbf{Centrosymmetry}: IC(k,t) = IC(K+1-k, T+1-t)$$

Further, for Models 1 and 2:

$$\textbf{Information-free cells}: IC\left(\tfrac{K+1}{2}, 1\right) = IC\left(\tfrac{K+1}{2}, T\right) = 1$$



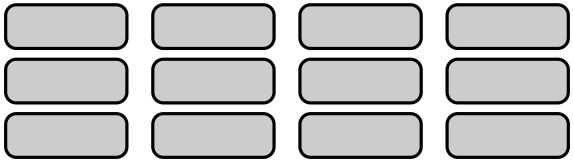
Information content of cells: theoretical results

For Models 1, 2, and 3 we get the following property:

$$\textbf{Centrosymmetry}: IC(k,t) = IC(K+1-k, T+1-t)$$

Further, for Models 1 and 2:

$$\textbf{Information-free cells}: IC\left(\tfrac{K+1}{2}, 1\right) = IC\left(\tfrac{K+1}{2}, T\right) = 1$$



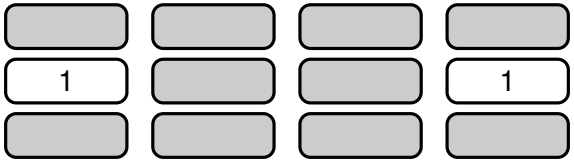
Information content of cells: theoretical results

For Models 1, 2, and 3 we get the following property:

$$\textbf{Centrosymmetry}: IC(k,t) = IC(K+1-k, T+1-t)$$

Further, for Models 1 and 2:

$$\textbf{Information-free cells}: IC\left(\tfrac{K+1}{2}, 1\right) = IC\left(\tfrac{K+1}{2}, T\right) = 1$$



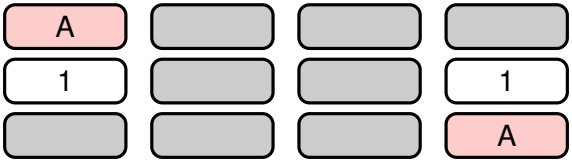
Information content of cells: theoretical results

For Models 1, 2, and 3 we get the following property:

$$\textbf{Centrosymmetry}: IC(k,t) = IC(K+1-k, T+1-t)$$

Further, for Models 1 and 2:

$$\textbf{Information-free cells}: IC\left(\tfrac{K+1}{2}, 1\right) = IC\left(\tfrac{K+1}{2}, T\right) = 1$$



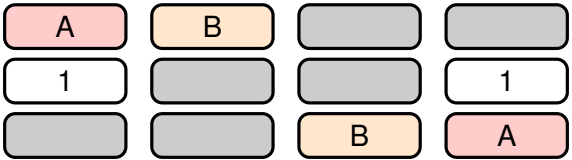
Information content of cells: theoretical results

For Models 1, 2, and 3 we get the following property:

$$\textbf{Centrosymmetry}: IC(k,t) = IC(K+1-k, T+1-t)$$

Further, for Models 1 and 2:

$$\textbf{Information-free cells}: IC\left(\tfrac{K+1}{2}, 1\right) = IC\left(\tfrac{K+1}{2}, T\right) = 1$$



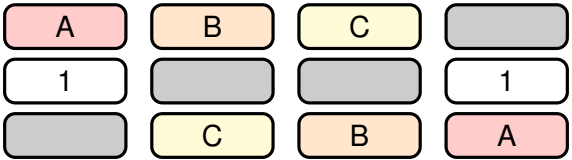
Information content of cells: theoretical results

For Models 1, 2, and 3 we get the following property:

$$\textbf{Centrosymmetry}: IC(k,t) = IC(K+1-k, T+1-t)$$

Further, for Models 1 and 2:

$$\textbf{Information-free cells}: IC\left(\tfrac{K+1}{2}, 1\right) = IC\left(\tfrac{K+1}{2}, T\right) = 1$$



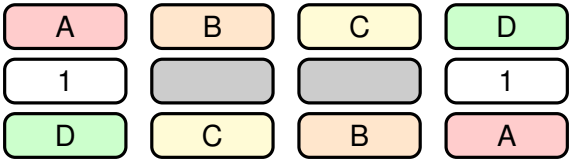
Information content of cells: theoretical results

For Models 1, 2, and 3 we get the following property:

$$\textbf{Centrosymmetry}: IC(k,t) = IC(K+1-k, T+1-t)$$

Further, for Models 1 and 2:

$$\textbf{Information-free cells}: IC\left(\tfrac{K+1}{2}, 1\right) = IC\left(\tfrac{K+1}{2}, T\right) = 1$$



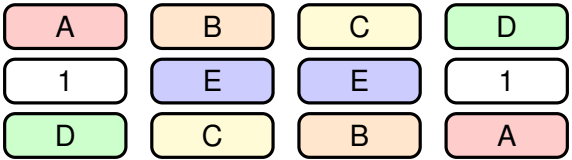
Information content of cells: theoretical results

For Models 1, 2, and 3 we get the following property:

$$\textbf{Centrosymmetry}: IC(k,t) = IC(K+1-k, T+1-t)$$

Further, for Models 1 and 2:

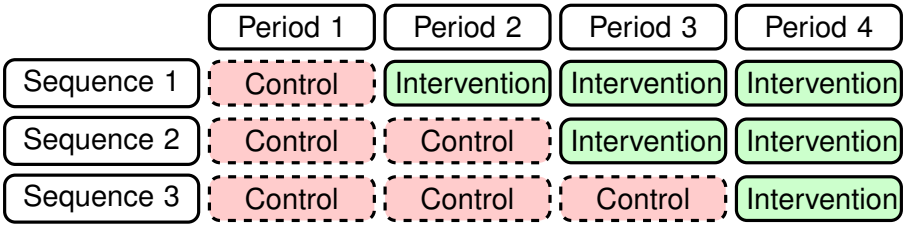
$$\textbf{Information-free cells}: IC\left(\tfrac{K+1}{2}, 1\right) = IC\left(\tfrac{K+1}{2}, T\right) = 1$$



Side note: On the symmetry of the stepped wedge

The stepped wedge possesses a particular symmetry:

- If time were reversed and treatment labels interchanged, we would get the same design.



Side note: On the symmetry of the stepped wedge

The stepped wedge possesses a particular symmetry:

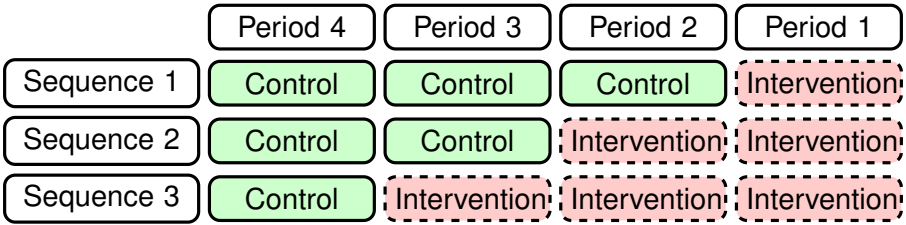
- If time were reversed and treatment labels interchanged, we would get the same design.

	Period 4	Period 3	Period 2	Period 1
Sequence 1	Intervention	Intervention	Intervention	Control
Sequence 2	Intervention	Intervention	Control	Control
Sequence 3	Intervention	Control	Control	Control

Side note: On the symmetry of the stepped wedge

The stepped wedge possesses a particular symmetry:

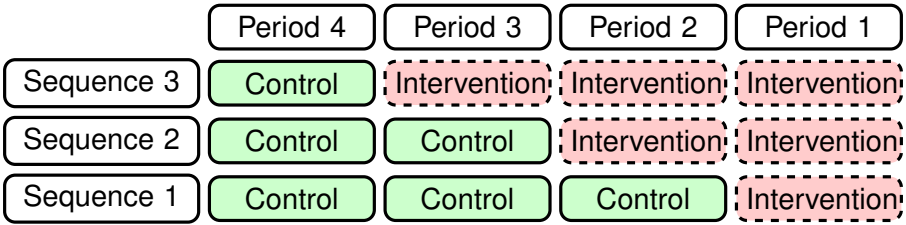
- If time were reversed and treatment labels interchanged, we would get the same design.



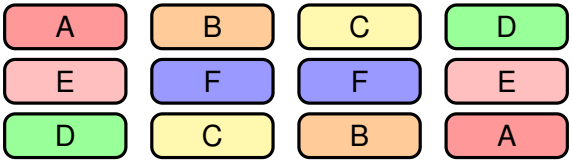
Side note: On the symmetry of the stepped wedge

The stepped wedge possesses a particular symmetry:

- If time were reversed and treatment labels interchanged, we would get the same design.

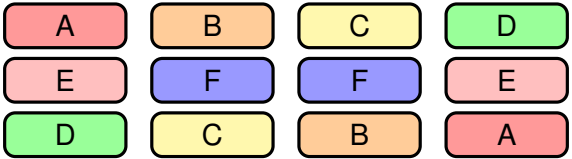


Side note: On the symmetry of the stepped wedge



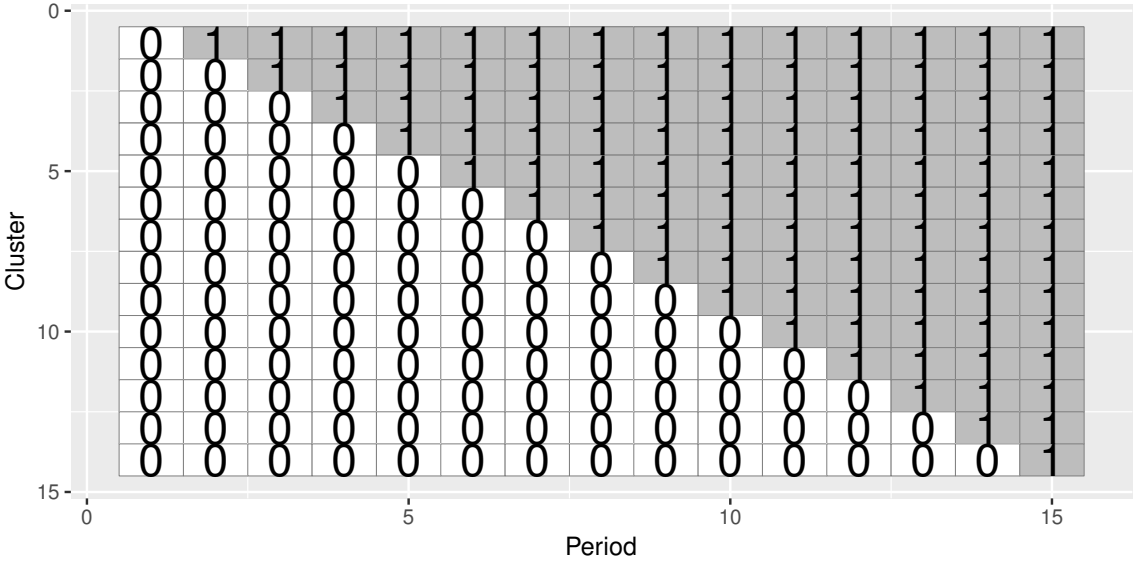
- We've shown that optimal estimators for the treatment effect must weight data from cluster-period cells in a *centrosymmetric* fashion.

Side note: On the symmetry of the stepped wedge

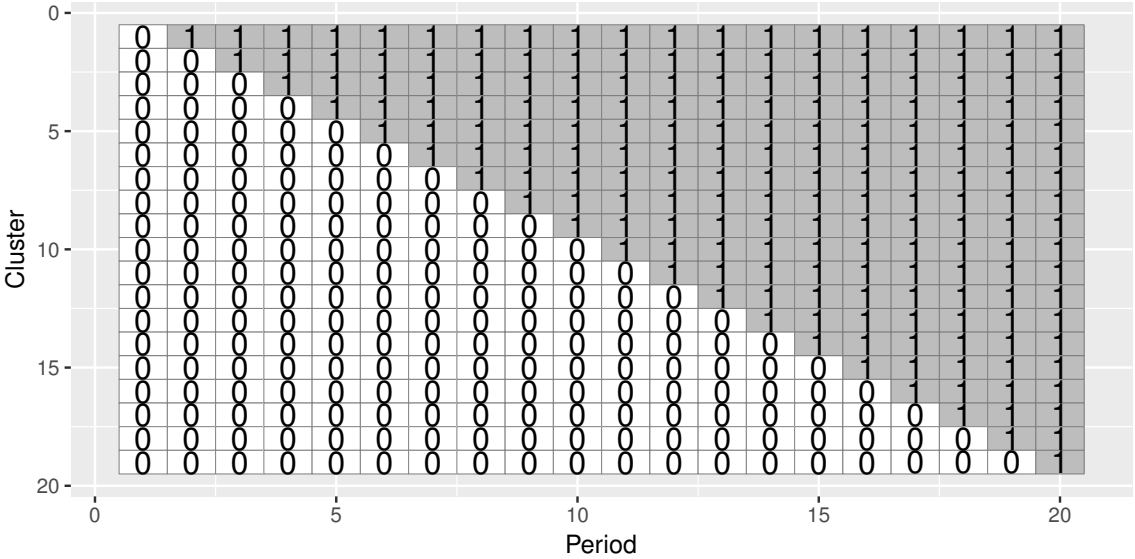


- We've shown that optimal estimators for the treatment effect must weight data from cluster-period cells in a *centrosymmetric* fashion.
  - And any estimators that aren't symmetric can be symmetrized!

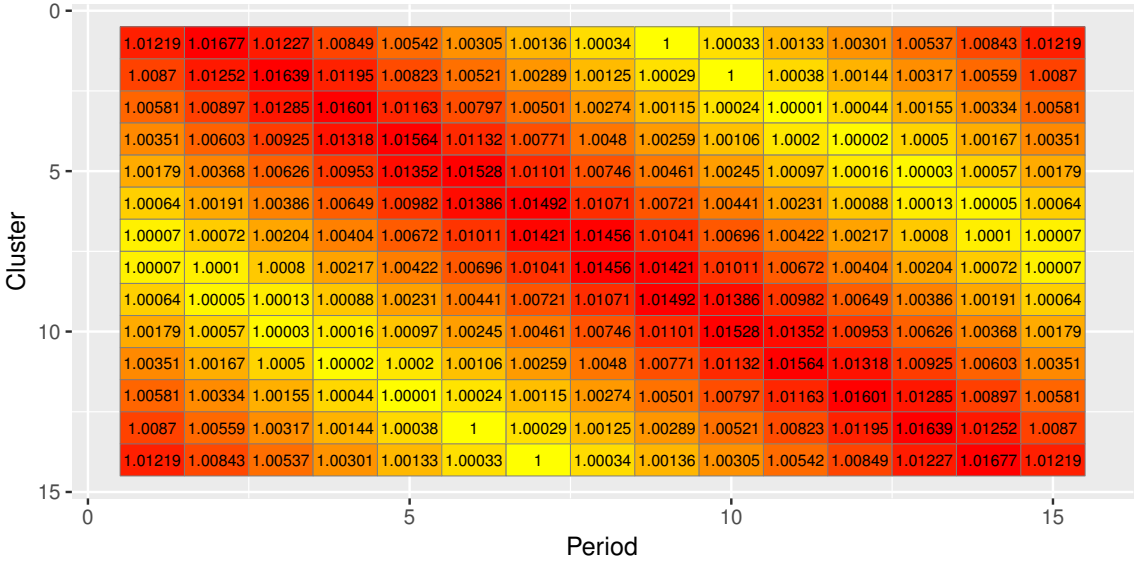
Design matrix: $T = 15$



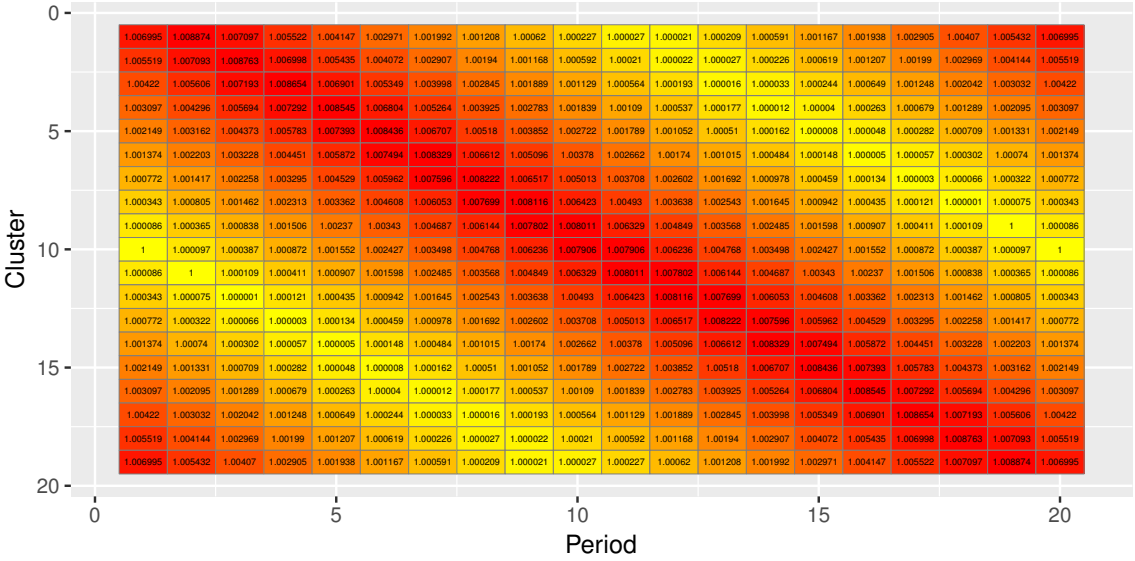
Design matrix: $T = 20$



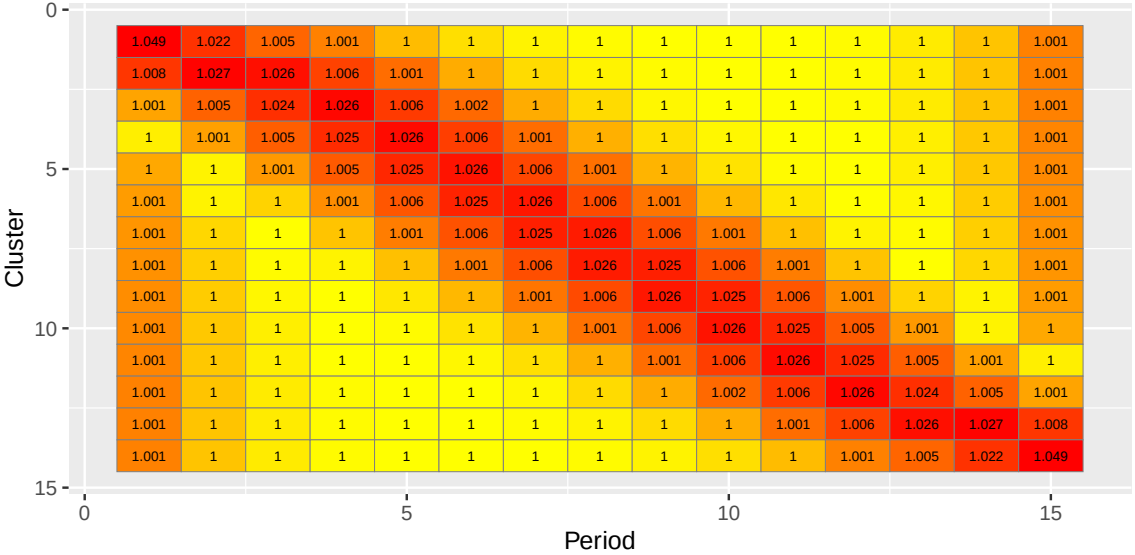
Information content of cells, Model 1: $\rho = 0.05$, $m = 100$



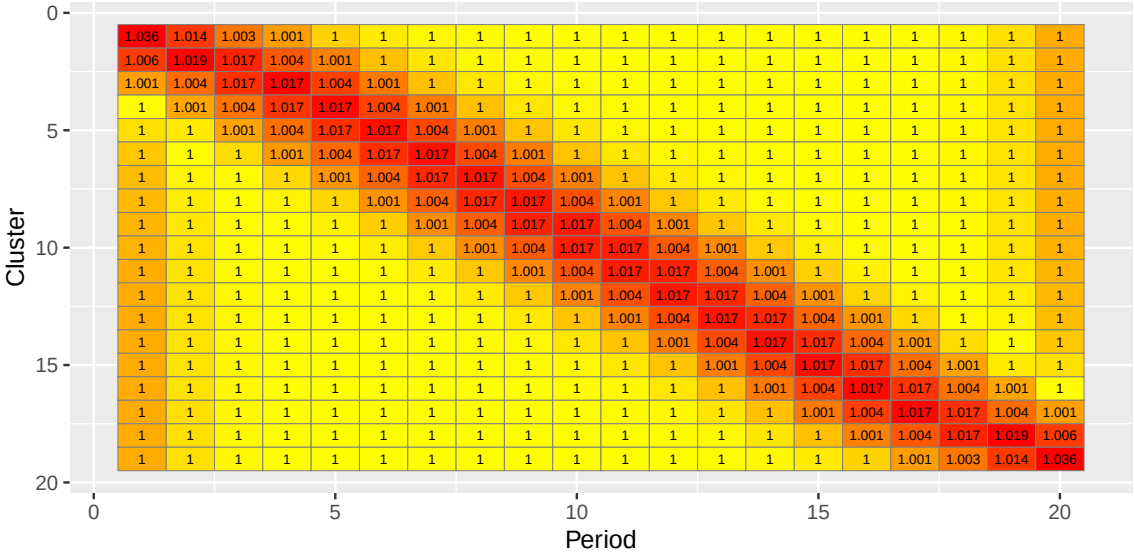
Information content of cells, Model 1: $\rho = 0.05$, $m = 100$



Information content of cells, Model 3: $\rho = 0.05$, r = 0.95, $m = 100$



Information content of cells, Model 3: $\rho = 0.05$, r = 0.95, $m = 100$



Pointing the way to incomplete designs

- Not all cells in the stepped wedge are created equal!
- Do we *really* need to include cells with low information content?
  - Can we exclude the cells that aren't contributing much?
- How much is this going to impact the power of the study to detect an effect?

Considering a real trial in Australian rehabilitation units[1]

- Assess the impact of individual education in addition to usual care on fall rates in 8 hospital rehabilitation units
- 5-period stepped wedge design:

	Period 1	Period 2	Period 3	Period 4	Period 5
Sequence 1	Control	Intervention	Intervention	Intervention	Intervention
Sequence 2	Control	Control	Intervention	Intervention	Intervention
Sequence 3	Control	Control	Control	Intervention	Intervention
Sequence 4	Control	Control	Control	Control	Intervention

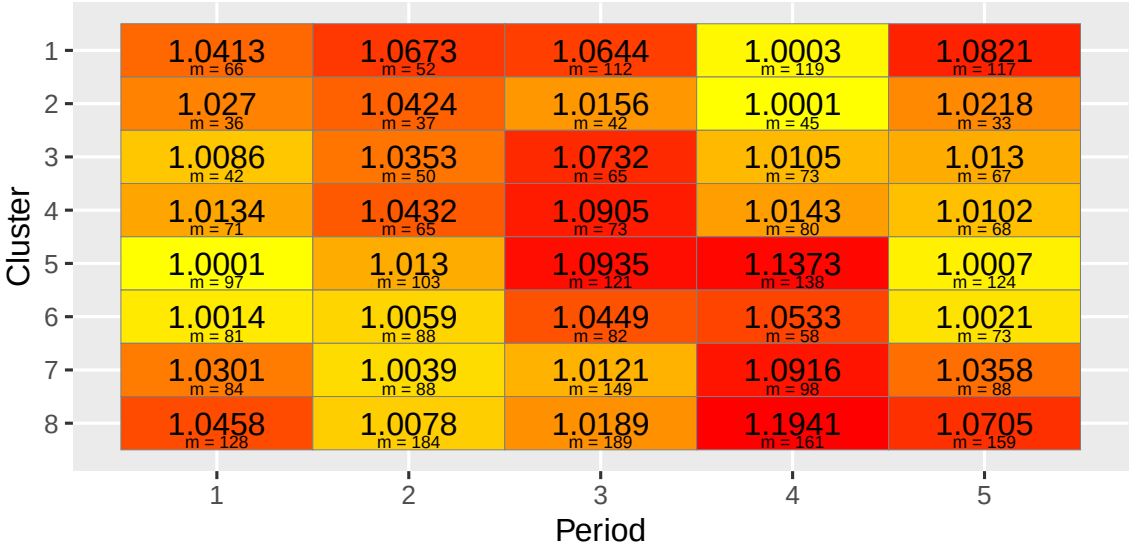
[1] Hill et al. The Lancet, 2015.

Differing numbers of patients in each cluster-period

The trial did not contain equal numbers of patients in each unit in each period!

	1	2	3	4	5
Cluster 1	66	52	112	119	117
Cluster 2	36	37	42	45	33
Cluster 3	42	50	65	73	67
Cluster 4	71	65	73	80	68
Cluster 5	97	103	121	138	124
Cluster 6	81	88	82	58	73
Cluster 7	128	184	189	161	159
Cluster 8	84	88	149	98	88

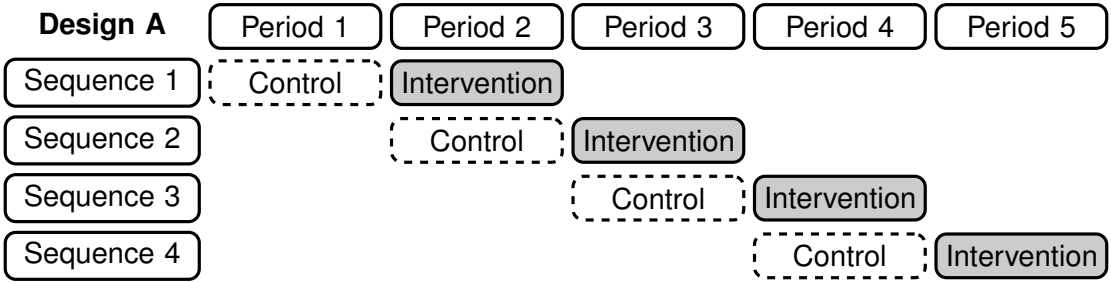
Information content of cluster-period cells in the rehab unit trial



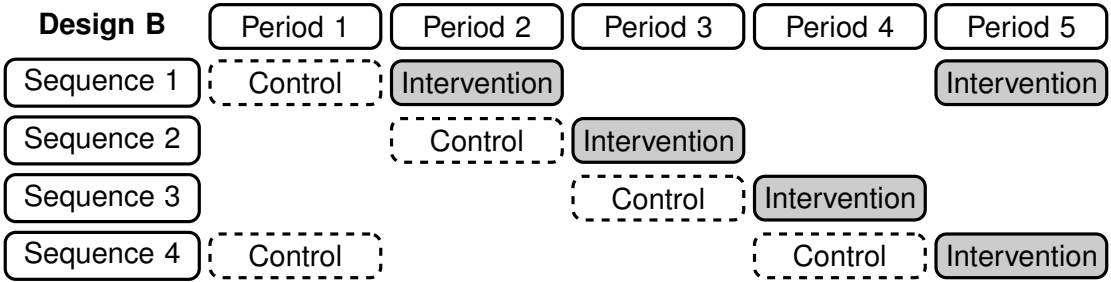
Do we need a complete design?

- Do we really need to recruit participants in each rehab centre in each period?
- Can we exclude some cluster-period cells?
  - $2^{20} - 1 = 1048575$ possible incomplete designs
- Instead of considering all of these, pick designs informed by information content pattern.

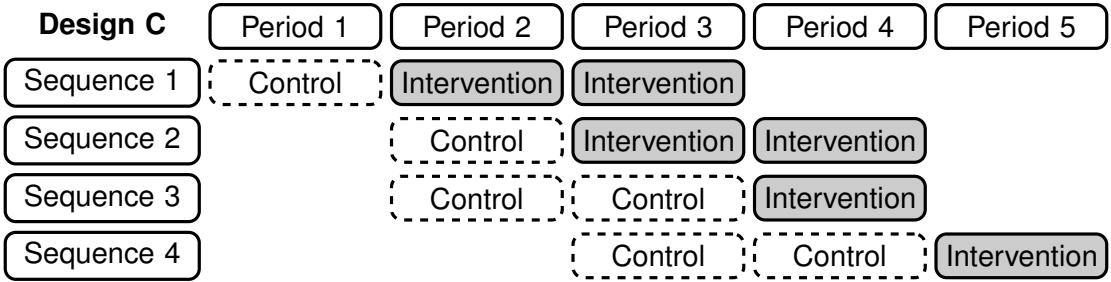
What about incomplete designs for this trial?



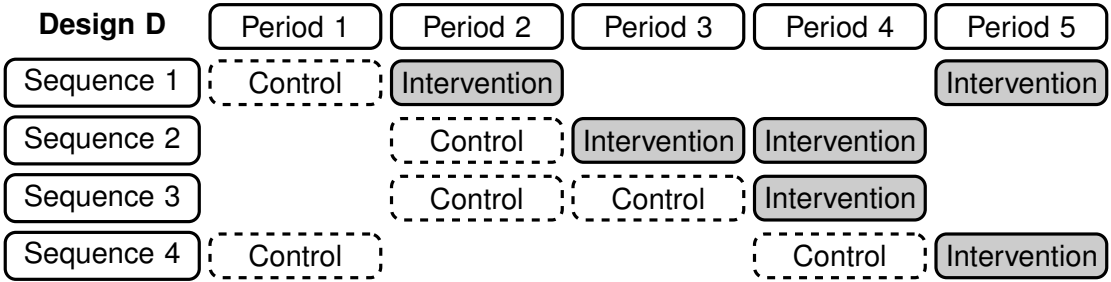
What about incomplete designs for this trial?



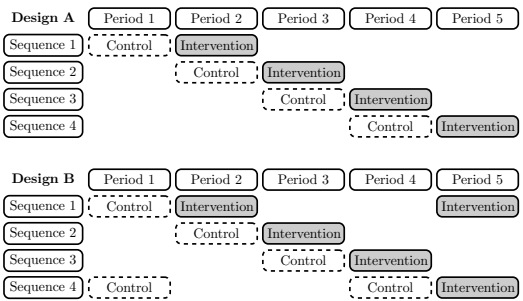
What about incomplete designs for this trial?



What about incomplete designs for this trial?

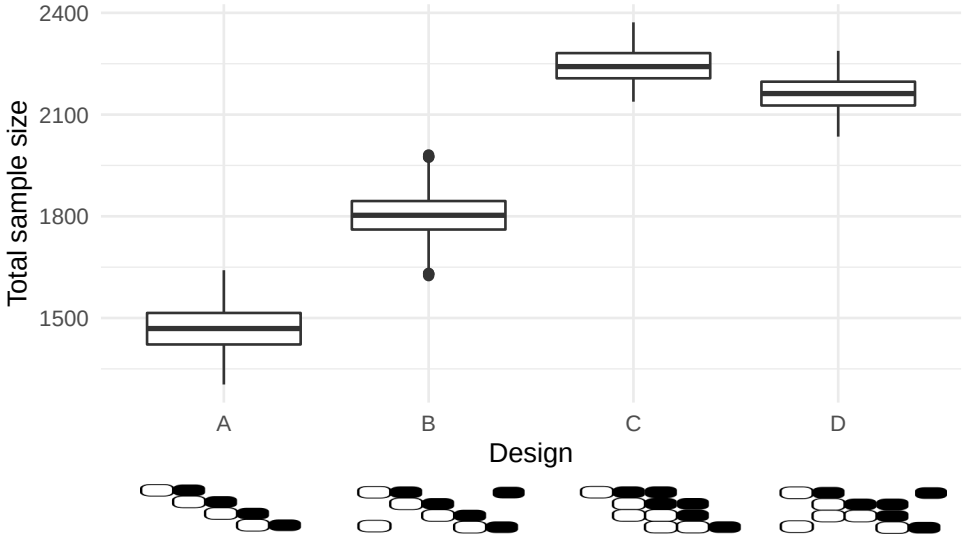


What about incomplete designs for this trial?

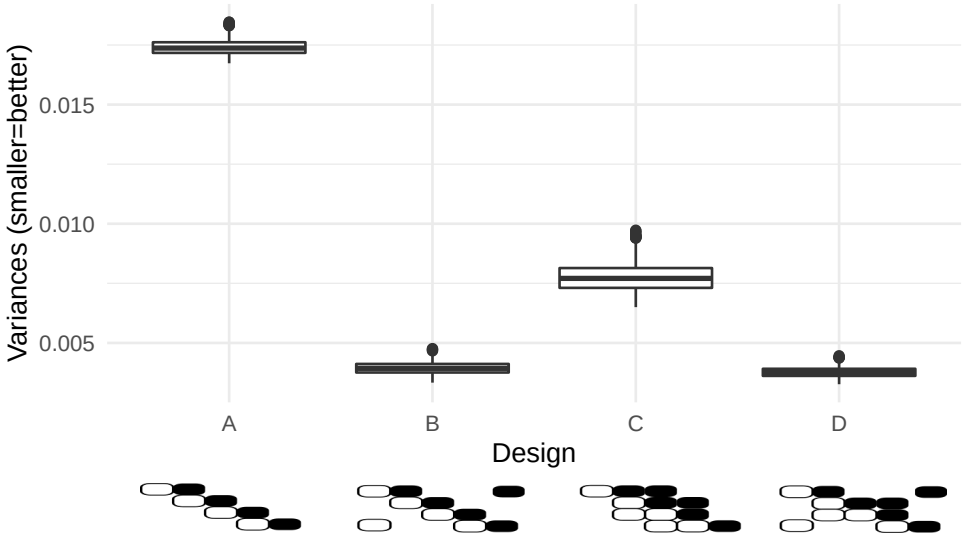


- Calculate $var(\hat{\theta})$ for each allocation of the 8 clusters to the 4 sequences (with 2 clusters per sequence)
  - 2520 possible allocations of clusters to sequences.

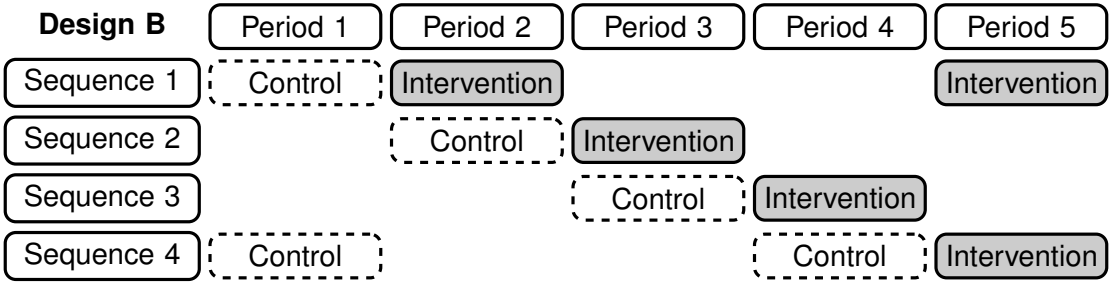
Total sample sizes of the incomplete Hill trials



The importance of “hot-spots” in the Hill et al. trial



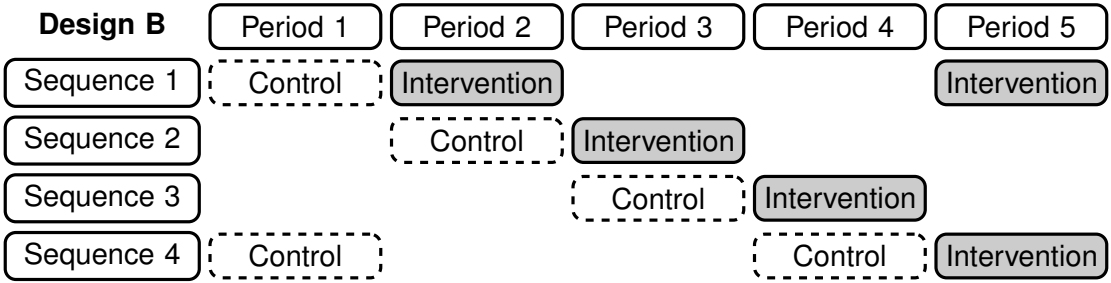
The winning design



- This design: 91% power[2]
  - Complete design: 94% power[3].

[2] To detect an effect of size 0.25 with 2-sided significance level of 0.05; ICC = 0.14; 90 participants per cluster-period.

The winning design



- This design: 91% power[2]
  - Complete design: 94% power[3].
- Are the "corner" measurements practical?

[2]To detect an effect of size 0.25 with 2-sided significance level of 0.05; ICC = 0.14; 90 participants per cluster-period.

What have we learned about the stepped wedge design?

- Periods near the treatment cross-over tend to be most valuable...
  - But the “hot corners” can add a lot of information (necessary to account for time effects)
  - Pattern of information content depends on the within-cluster correlation structure.
- Logistical importance vs. statistical value of cells?

Future work: development of “optimal” incomplete designs.

- How can we use the information content to inform selection of incomplete designs in a principled way? (Ehsan Rezaei)
- Investigating the “staircase designs”: when are these an efficient design choice? (Kelsey Grantham)

Questions? Ask now, or get in touch:

jessica.kasza@monash.edu

If you want to read more

...about within-cluster correlation structures:

- Grantham, Kasza, Heritier, Hemming, Forbes. *Accounting for a decaying correlation structure in cluster randomized trials with continuous recruitment*. **Statistics in Medicine**. 2019.
- Kasza, Hemming, Hooper, Matthews, Forbes. *Impact of non-uniform correlation structure on sample size and power in multiple-period cluster randomised trials*. **Statistical Methods in Medical Research**. 2019.

... about the information content of stepped wedge designs:

- Kasza, Forbes. *Information content of cluster-period cells in stepped wedge trials*. **Biometrics**. 2019.
- Kasza, Taljaard, Forbes. *Information content of stepped-wedge designs when treatment effect heterogeneity and/or implementation periods are present*. **Statistics in Medicine**. 2019.
- Kasza, Bowden, Forbes. *Information content of stepped wedge designs with unequal cluster-period sizes in linear mixed models: Informing incomplete desgins*. **Statistics in Medicine**. 2020.

... about the centrosymmetry of the stepped wedge:

- Bowden, Forbes, Kasza. *On the centrosymmetry of treatment effect estimators for stepped wedge and related cluster randomized trial designs*. **Statistics & Probability Letters**. 2021.