Information content of cluster-period cells in stepped wedge trials

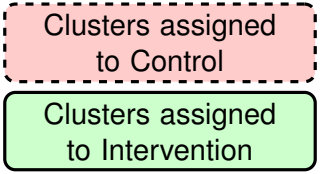
Jessica Kasza: jessica.kasza@monash.edu

Andrew Forbes: andrew.forbes@monash.edu

MONASH University

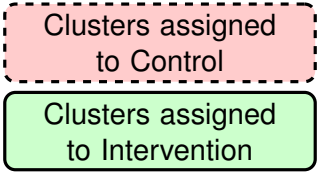
International Biometric Society Journal Club, February 2020

The standard cluster randomised trial



- *Clusters (groups)* of participants assigned to treatments. (Why?)
- Clusters could be hospitals, intensive care units, schools, neighbourhoods...

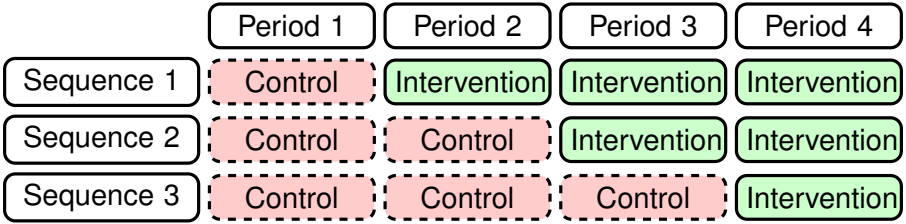
The standard cluster randomised trial



- *Clusters (groups)* of participants assigned to treatments. (Why?)
- Clusters could be hospitals, intensive care units, schools, neighbourhoods...

Can increase efficiency by considering *longitudinal (multiple period)* cluster randomised trials.

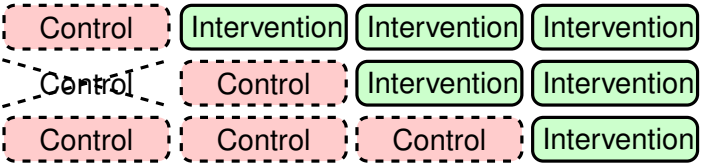
The stepped wedge cluster randomised trial design



- Stepped wedge designs can be useful when all clusters need to receive the intervention, or the intervention is going to be rolled out anyway.
- K clusters are randomised to $T - 1$ sequences; $K \times T$ cluster-period cells;
- m participants in each cluster in each period.

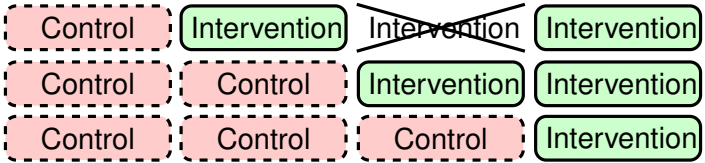
Each cluster-period pair is a **cell** of the design.

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 cluster $k = 1, \ldots, K$, in period $t = 1, \ldots, T$:

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

$$\mathbf{CP}_k = (CP_{k1}, \ldots, CP_{kT}) \sim N_T(0, \mathbf{V}_{CP})$$

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 cluster $k = 1, \ldots, K$, in period $t = 1, \ldots, T$:

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

$$\mathbf{CP}_k = (CP_{k1}, \ldots, CP_{kT}) \sim N_T(0, \mathbf{V}_{CP})$$

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

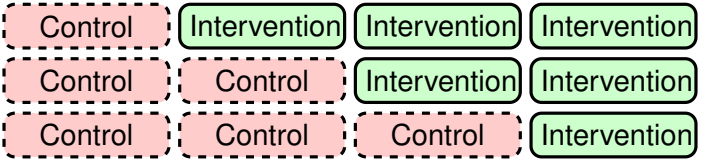
θ is the treatment effect: this is what we want to estimate.

- $\hat{\theta}$ the weighted 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?

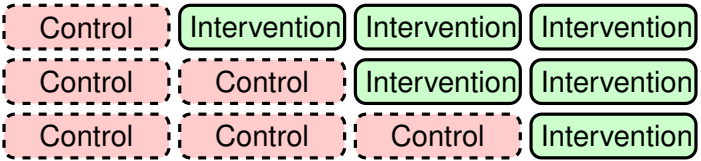
Information content of each cell

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

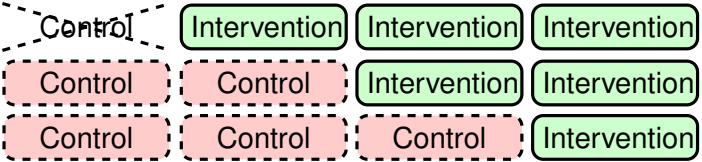


Information content of each cell

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

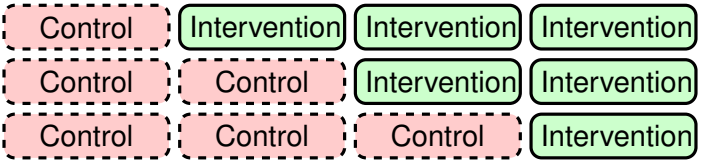


Calculate $var(\hat{\theta})_{[kt]}$ from the incomplete design, omitting period t of cluster k:

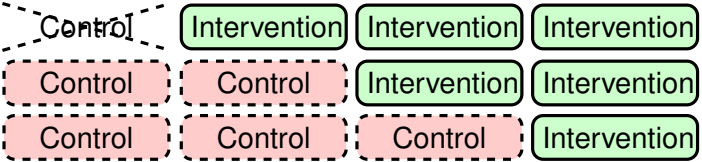


Information content of each cell

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



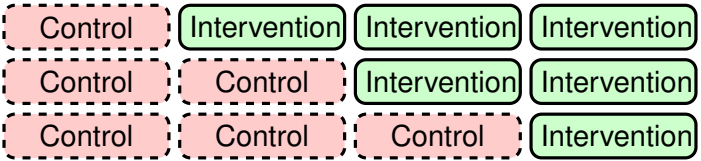
Calculate $var(\hat{\theta})_{[kt]}$ from the incomplete design, omitting period t of cluster k:



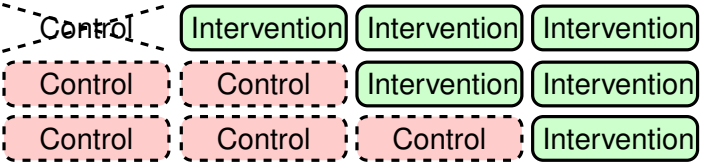
Information content of cell (k, t)**:** $IC(k,t) = var(\hat{\theta})_{[kt]}/var(\hat{\theta})$

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: theoretical results

Can obtain a closed-form expression for $IC(k,t) = var(\hat{\theta})_{[kt]}/var(\hat{\theta})$ for certain models[1]

- Depends on the **within-cluster correlation structure**.

$$Y_{kti} = \mu + \beta_t + X_{kt}\theta + CP_{kt} + \epsilon_{kti}, \quad \epsilon_{kti} \sim N(0, \sigma^2_\epsilon)$$
$$\mathbf{CP}_k = (CP_{k1}, \ldots, CP_{kT}) \sim N_T(0, \mathbf{V}_{CP})$$

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

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

[1] A closed form expression for $IC(k,t)$ is available whenever the inverse of the covariance matrix of observations from a cluster has a closed form.

Within-cluster correlation structure

Model 1: Hussey and Hughes (2007)

$$CP_{kt} = CP_{ks} = CP_k \sim N(0, \tau^2)$$

$$\rho = corr(Y_{kti}, Y_{ksj}) = \frac{\tau^2}{\tau^2 + \sigma_\epsilon^2}$$

- Correlation between the outcomes of any pair of participants is identical.

Period	1	2	3	4
$corr(Y_{k1i}, Y_{ksj})$	ρ	ρ	ρ	ρ

Within-cluster correlation structure

Model 2: "Constant between-period correlation model"

$$\mathbf{CP}_k \sim N_T\left(0, \tau^2\left[r_0 J_T + (1 - r_0) I_T\right]\right)$$

$$\rho = corr(Y_{kti}, Y_{ktj}) = \frac{\tau^2}{\tau^2 + \sigma_\epsilon^2}, \quad corr(Y_{kti}, Y_{ksj}) = r_0 \frac{\tau^2}{\tau^2 + \sigma_\epsilon^2} = r_0 \rho$$

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

Period	1	2	3	4
$corr(Y_{k1i}, Y_{ksj})$	ρ	$r_0\rho$	$r_0\rho$	$r_0\rho$

Within-cluster correlation structure

Model 3: "Exponential decay model"

$$\mathbf{CP}_k \sim N_T\left(0, \tau^2 R\right), \quad R[t,s] = r^{|t-s|}$$

$$\rho = corr(Y_{kti}, Y_{ktj}) = \frac{\tau^2}{\tau^2 + \sigma_\epsilon^2}, \quad corr(Y_{kti}, Y_{ksj}) = r^{|t-s|}\frac{\tau^2}{\tau^2 + \sigma_\epsilon^2} = r^{|t-s|}\rho$$

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

Period	1	2	3	4
$corr(Y_{k1i}, Y_{ksj})$	ρ	$r\rho$	$r^2\rho$	$r^3\rho$

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

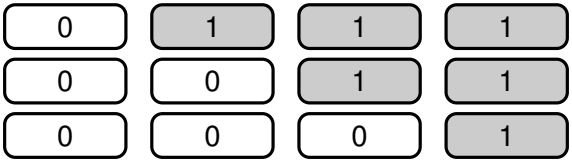
Information content of cells: theoretical results

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:

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



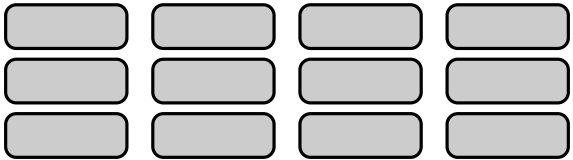
Information content of cells: theoretical results

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:

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



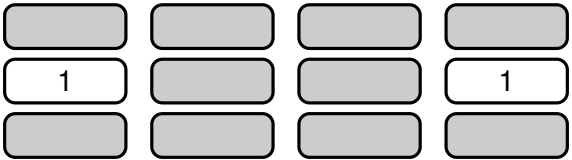
Information content of cells: theoretical results

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

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

Further, for Models 1 and 2:

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



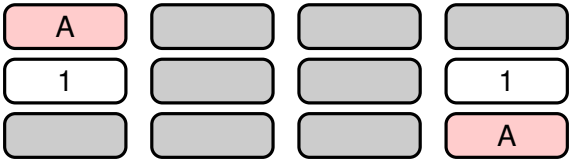
Information content of cells: theoretical results

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:

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



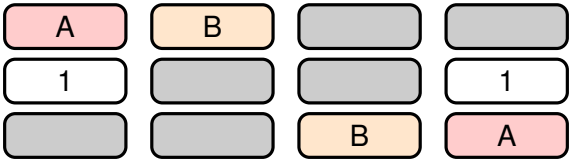
Information content of cells: theoretical results

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:

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



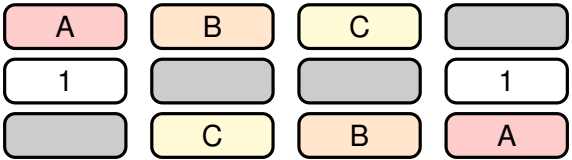
Information content of cells: theoretical results

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:

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



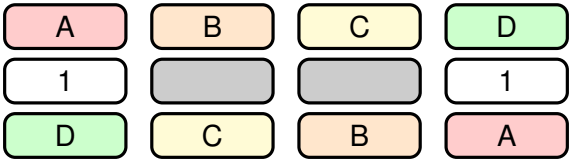
Information content of cells: theoretical results

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:

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



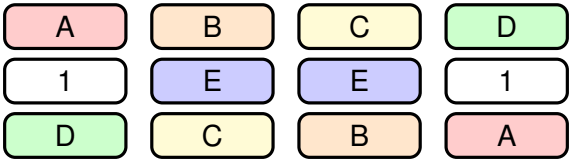
Information content of cells: theoretical results

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

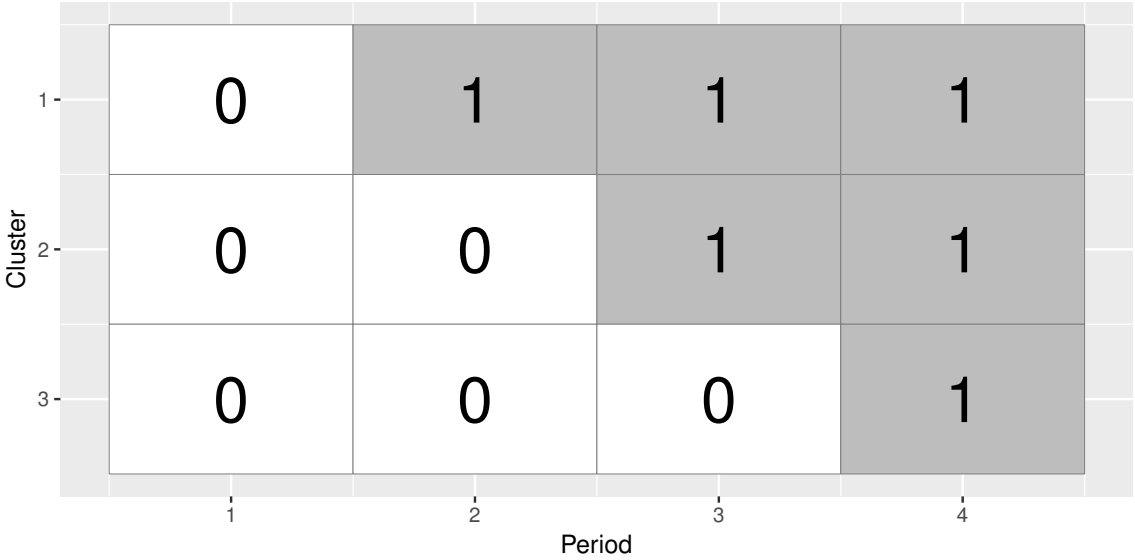
$$\textbf{Centrosymmetry}: IC(k,t) = IC(K+1-k, T+1-t)$$

Further, for Models 1 and 2:

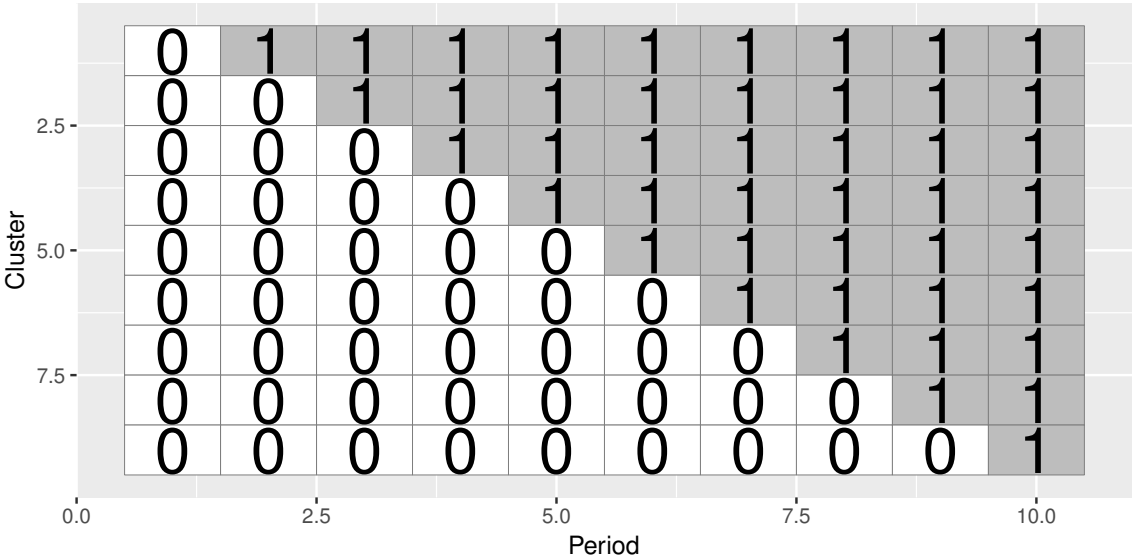
$$\textbf{Information-free cells}: IC\left(\tfrac{K+1}{2}, 1\right) = IC\left(\tfrac{K+1}{2}, T\right) = 1$$



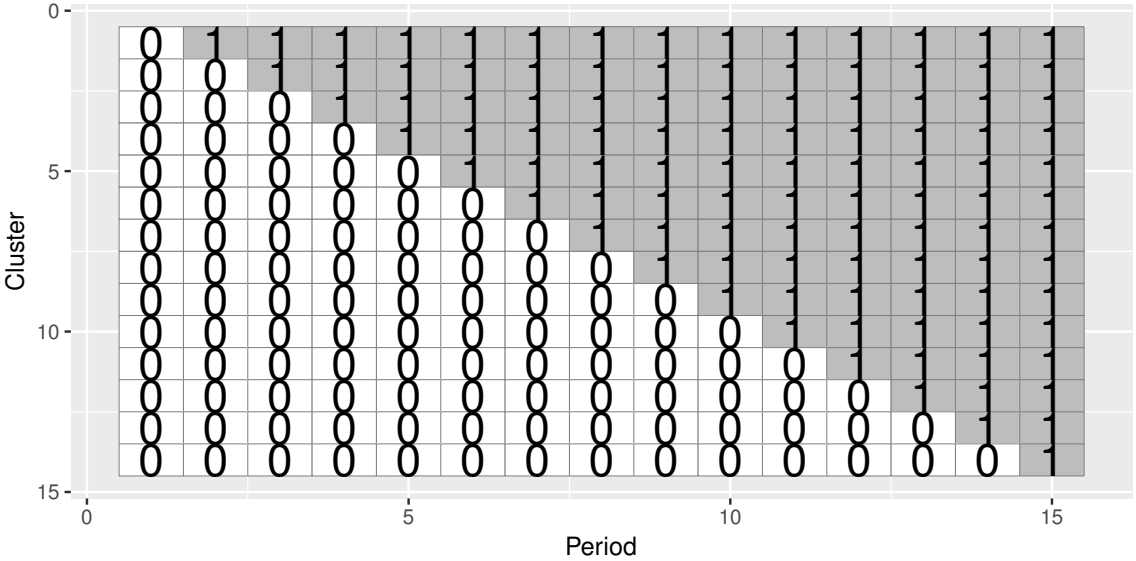
Design matrix: $T = 4$



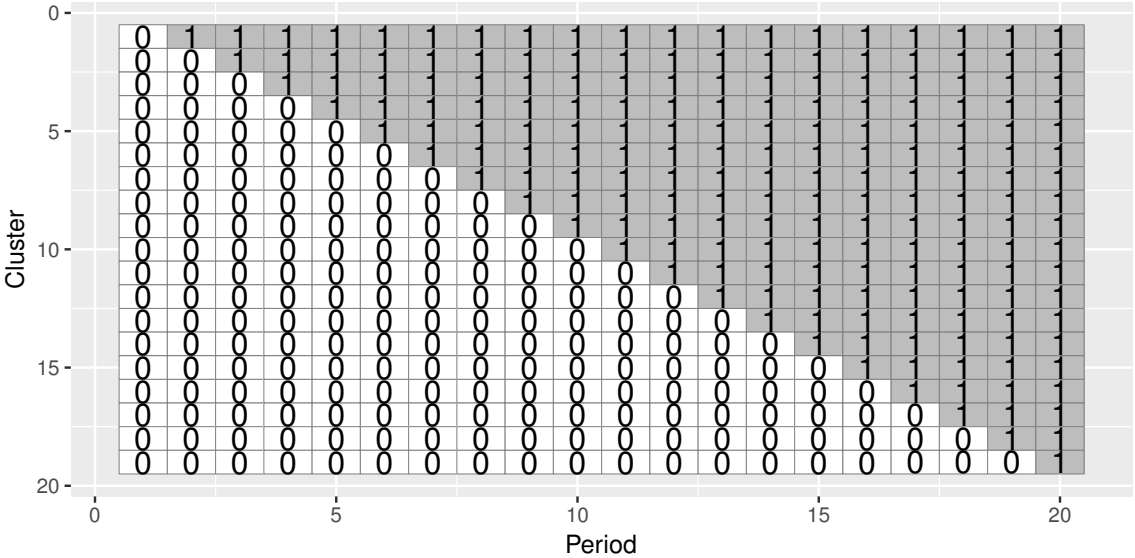
Design matrix: $T = 10$



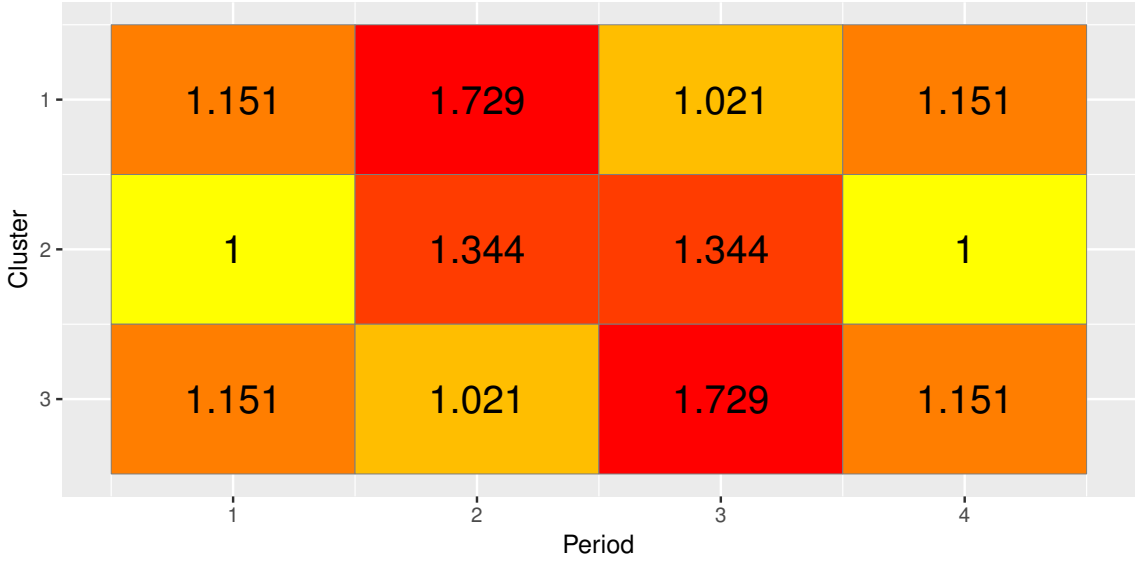
Design matrix: $T = 15$



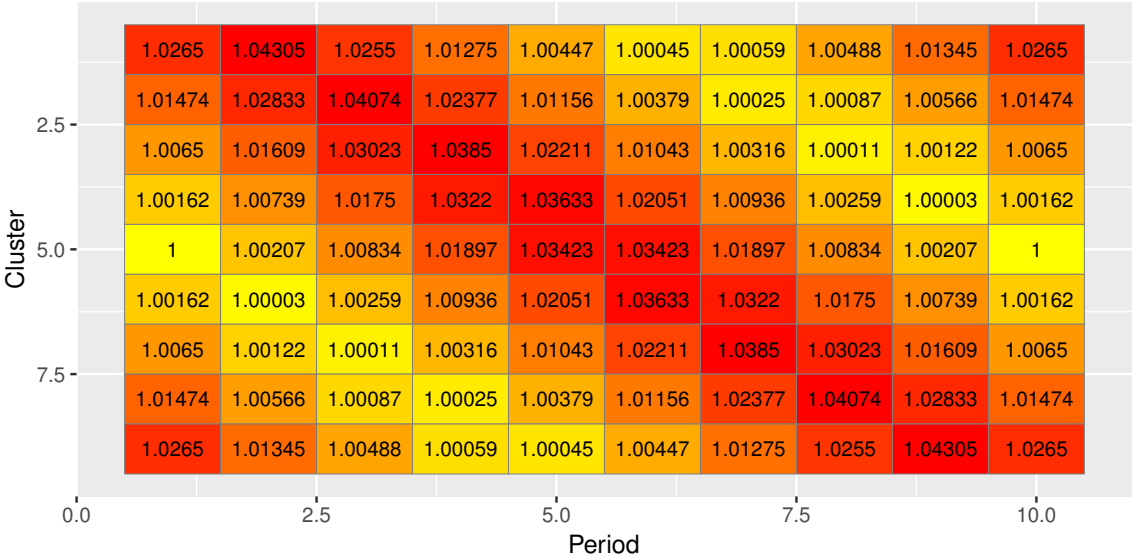
Design matrix: $T = 20$



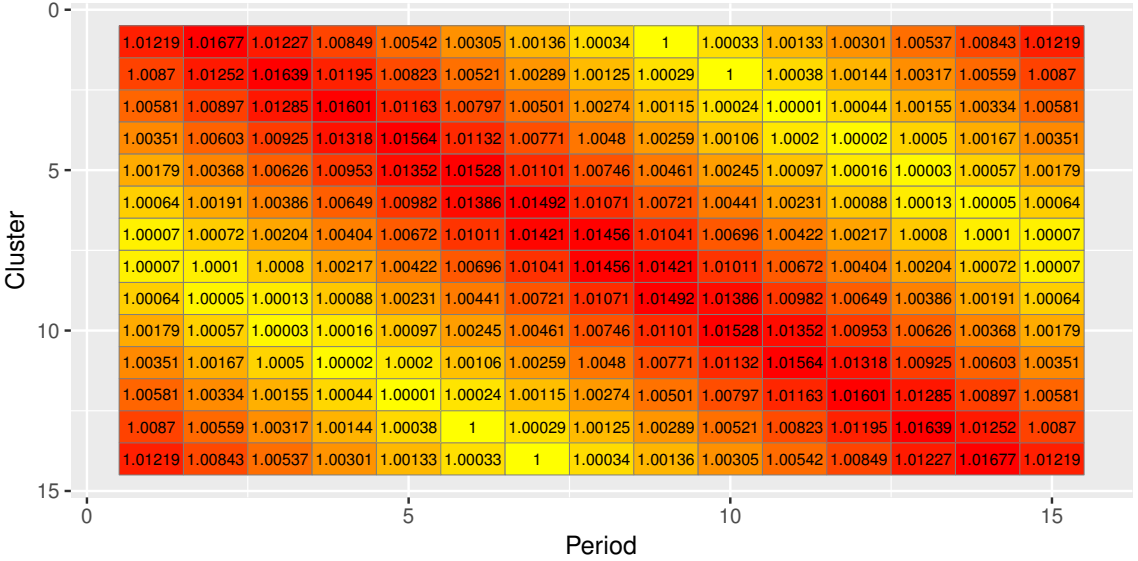
Information content of cells, Model 1: $\rho = 0.05$, $m = 100$



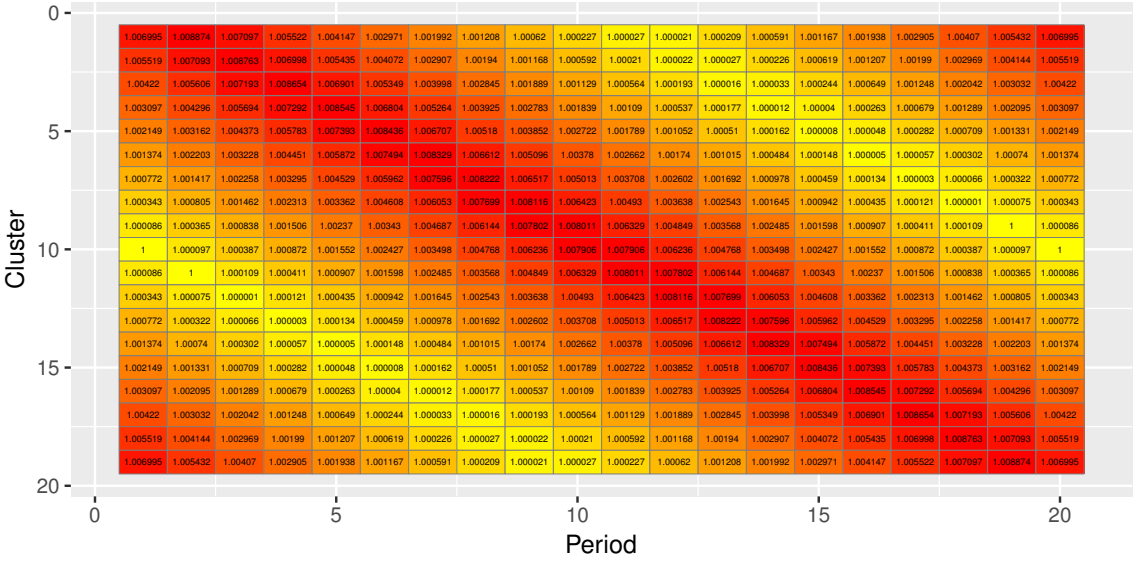
Information content of cells, Model 1: $\rho = 0.05$, $m = 100$



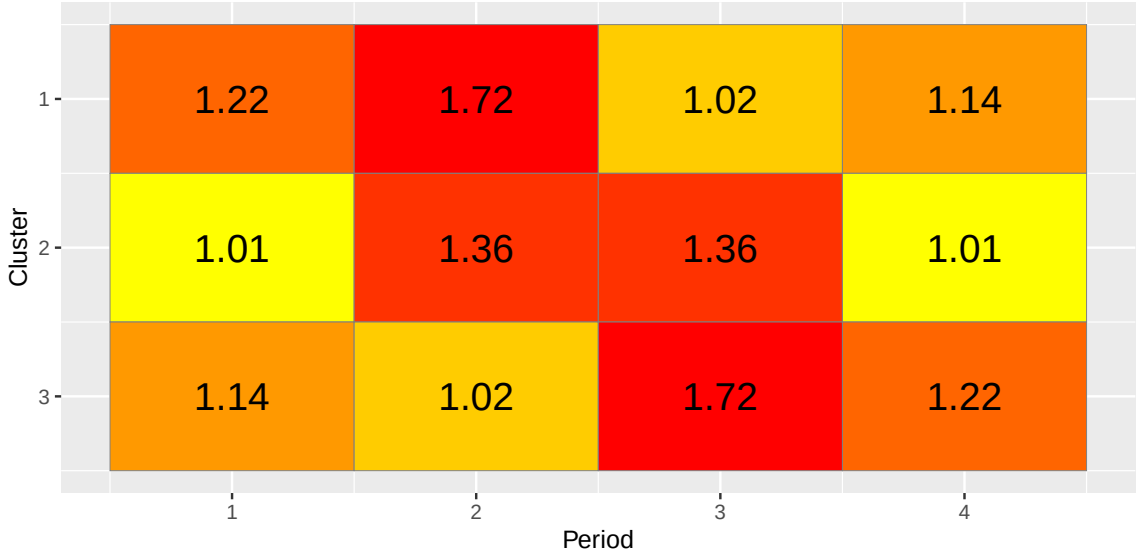
Information content of cells, Model 1: $\rho = 0.05$, $m = 100$



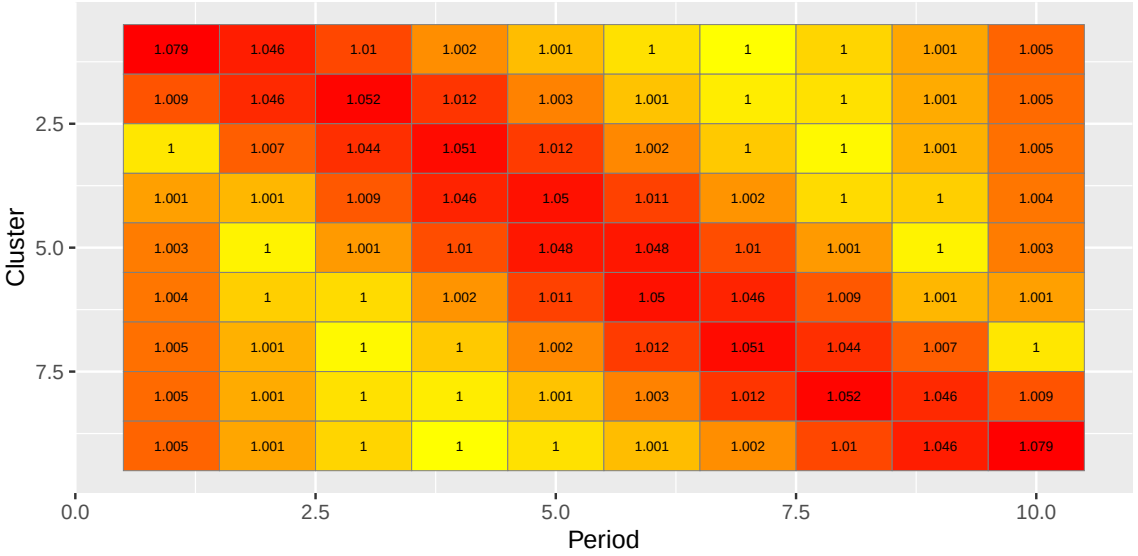
Information content of cells, Model 1: $\rho = 0.05$, $m = 100$



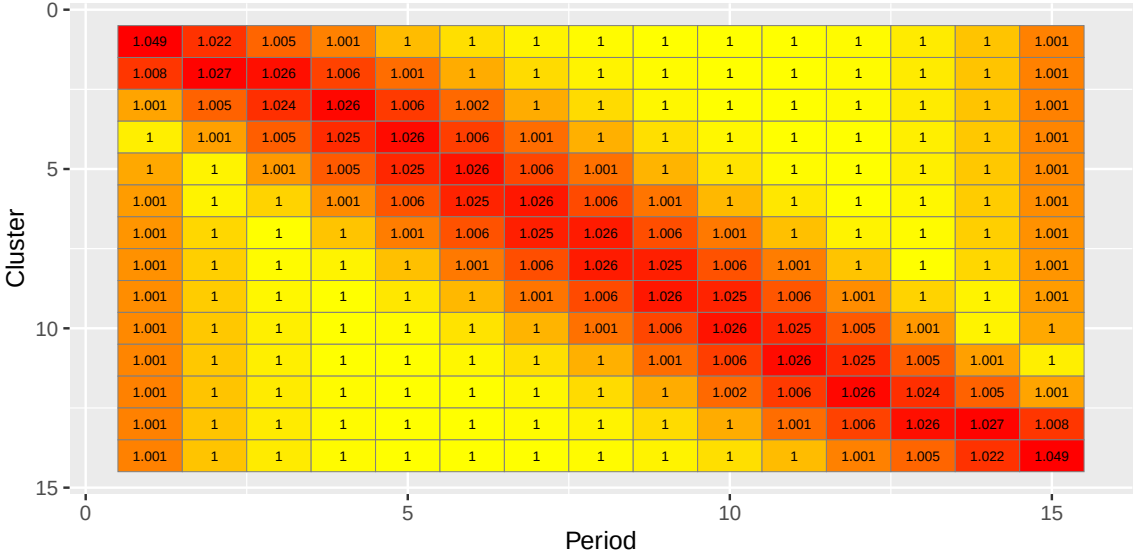
Information content of cells, Model 3: $\rho = 0.05$, r = 0.95, $m = 100$



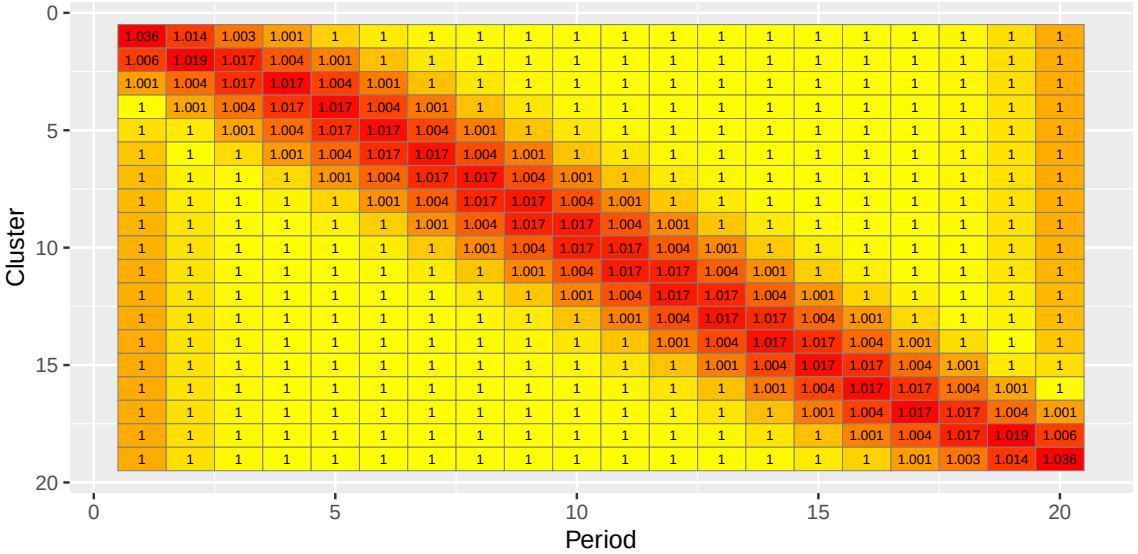
Information content of cells, Model 3: $\rho = 0.05$, r = 0.95, $m = 100$



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



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



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 vs. statistical value of cells?

Here we assumed a very simple situation. But what if....

- there are transition periods (i.e. periods missing by design)?
- there is treatment effect heterogeneity?
- clusters/cells are of different sizes?
- a different treatment effect estimator is considered?

Future work: development of “optimal” incomplete designs.

You can explore the information content of cells in your own cluster randomised trial at:

`https://monash-biostat.shinyapps.io/InformationContentofCells`