Contamination: How much can an individually randomised trial tolerate?

K Hemming; M Taljaard; A Forbes

Birmingham University; OHRI; Monash University

November 11, 2019

Background

- Cluster randomised trials:
  - **Sometimes used to avoid contamination**;
  - Result in a loss of statistical efficiency and are at increased risk of bias.
- Individually randomised trials:
  - Increased statistical efficiency;
  - But, if there is any contamination will attenuate treatment effect.
- But, the increased statistical efficiency of the individually randomised design *might* means the attenuated effect can be detected with a smaller sample size than under cluster randomization. [1]

[1] Slymen DJ, Hovell MF. Cluster versus individual randomization in adolescent tobacco and alcohol studies: illustrations for design decisions. Int J Epidemiol. 1997; 26(4):765-71. Torgerson DJ. Contamination in trials: is cluster randomisation the answer? BMJ. 2001 Feb 10;322(7282):355-7.

Overview

- Provide a general framework for determining the amount of contamination that can be tolerated in an individually randomised design before a larger sample size is required than a cluster randomised design.
- Consider a variety of designs:
  - Parallel arm designs with and without baselines;
  - Cluster crossover trials under a range of assumptions about the within-cluster correlation structure;
  - Unstratified and stratified individually randomised trials.
- We show that individually randomised trials can tolerate a surprisingly large amount of contamination.

*Note we are only considering one form of contamination: that of control arm contamination by the intervention.

*i*ndividually Randomised Controlled Trials (*i*RCTs)

The required **sample size** per arm for a trial at pre-specified power $1-\beta$ to detect a difference of δ, is n_I, where:

$$n_I = 2\sigma^2 \left[\frac{(z_{\alpha/2} + z_\beta)^2}{\delta^2} \right] \tag{1}$$

and where $z_{\alpha/2}$ is the critical value of the z-distribution with an area $\alpha/2$ in each tail.

Cluster randomised trials (CRTs)

The required **sample size** per arm to detect a difference of δ, in a CRT with an intra-cluster correlation (ICC) of ρ and cluster size m is: n_{CRT}, where:

$$n_{CRT} = 2\sigma^2 \left[\frac{(z_{\alpha/2} + z_{\beta})^2}{\delta^2} \right] [1 + (m-1)\rho] \tag{2}$$

$$= n_I \underbrace{[1 + (m-1)\rho]}_{DE} \tag{3}$$

where $DE = [1 + (m-1)\rho]$ is the Design Effect for clustering.

Individually randomised trials with contamination

The required sample size per arm for a trial to detect a difference of δ, where the **rate of contamination** is w, is n_I*, where:

$$n_I* = 2\sigma^2 \left[\frac{(z_{\alpha/2} + z_\beta)^2}{((1-w)\delta)^2} \right] = n_I \underbrace{(1-w)^{-2}}_{DE} \tag{4}$$

where $DE = [(1-w)^{-2}]$ is the **design effect for contamination**.

[1] Note rate of contamination represents the proportion of observations in the control condition that receive the full effect of the intervention condition

Derivation

For an individually randomised trial with contamination w, meaning $100w\%$ of control subjects get the intervention, then the expected mean in the control arm is:

$$(1-w)\mu_0 + w\mu_1 \tag{5}$$

where μ_0 and μ_1 are the means for the control and treatment conditions, when received. So the expected (attenuated) treatment difference is:

$$\underbrace{\mu_1}_{\textit{mean Tx}} - \underbrace{((1-w)\mu_0 + w\mu_1)}_{\textit{mean Tx'}} = (1-w)\delta \tag{6}$$

So the sample size per arm in an individually randomised trial with contamination is:

$$n_I* = 2\sigma^2 \left[\frac{(z_{\alpha/2} + z_\beta)^2}{((1-w)\delta)^2} \right] \tag{7}$$

Comparing design effects

How much contamination can be tolerated?

Determine the contamination rate at which an individually randomised trial with an attenuated treatment effect requires a larger sample size than a CRT.

When $n_I* > n_{CRT}$:

$$2\sigma^2 \left[\frac{(z_{\alpha/2} + z_\beta)^2}{((1-w)\delta)^2} \right] > 2\sigma^2 \left[\frac{(z_{\alpha/2} + z_\beta)^2}{\delta^2} \right] [1 + (m-1)\rho] \tag{8}$$

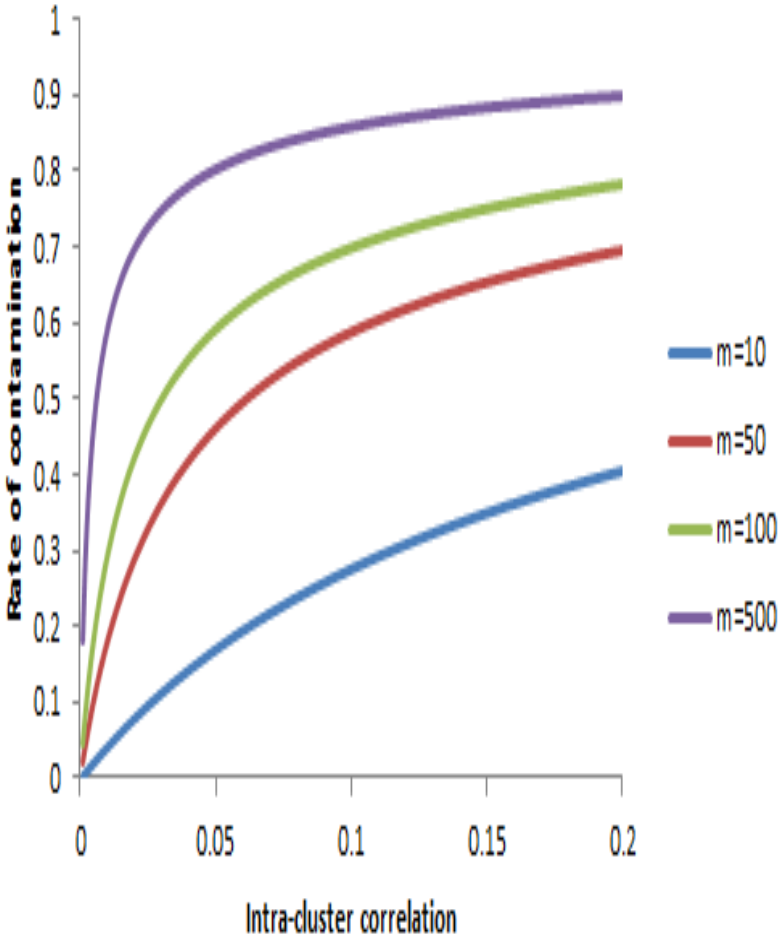
That is, when:

$$(1-w)^{-2} > [1 + (m-1)\rho] \tag{9}$$

so when:

$$w > 1 - [1 + (m-1)\rho]^{-1/2} \tag{10}$$

Critical values plot: CRT vs *i*RCT



Example (small cluster size)

For an ICC of 0.05 and $m = 10$, up to 20% contamination can be tolerated before the sample size exceeds that of a CRT.

Example (large cluster size)

For an ICC of 0.05 and $m = 500$, up to 80% contamination can be tolerated before the sample size exceeds that of a CRT.

Cluster randomised trial with baseline measures (CRT-B)

Parallel cluster trial with baseline measures

In a parallel cluster randomised trial with baseline measures all clusters are initially in the control condition and then (typically) half receive the intervention. Assume cross-sectional design.

The sample size per arm to detect δ, for within-period ICC of ρ and **cluster size** m **per period** and where η is the **cluster auto correlation (CAC)** is:

$$n_{CRT-B} = \underbrace{2\sigma^2 \left[\frac{(z_{\alpha/2} + z_{\beta})^2}{\delta^2}\right]}_{n_I} \underbrace{[1 + (m-1)\rho]}_{DE} 2(1 - r^2) \qquad (11)$$

where $r = \frac{m\rho\eta}{1+(m-1)\rho}$ is the cluster mean correlation.

[1] Teerenstra S, Eldridge S, Graff M, de Hoop E, Borm GF. A simple sample size formula for analysis of covariance in cluster randomized trials. Stat Med. 2012

Comparing design effects: CRT with baseline measures

How much contamination can be tolerated?

Determine the contamination rate at which an individually randomised trial with an attenuated treatment effect requires a larger sample size than a cluster randomised trial with baseline measures.

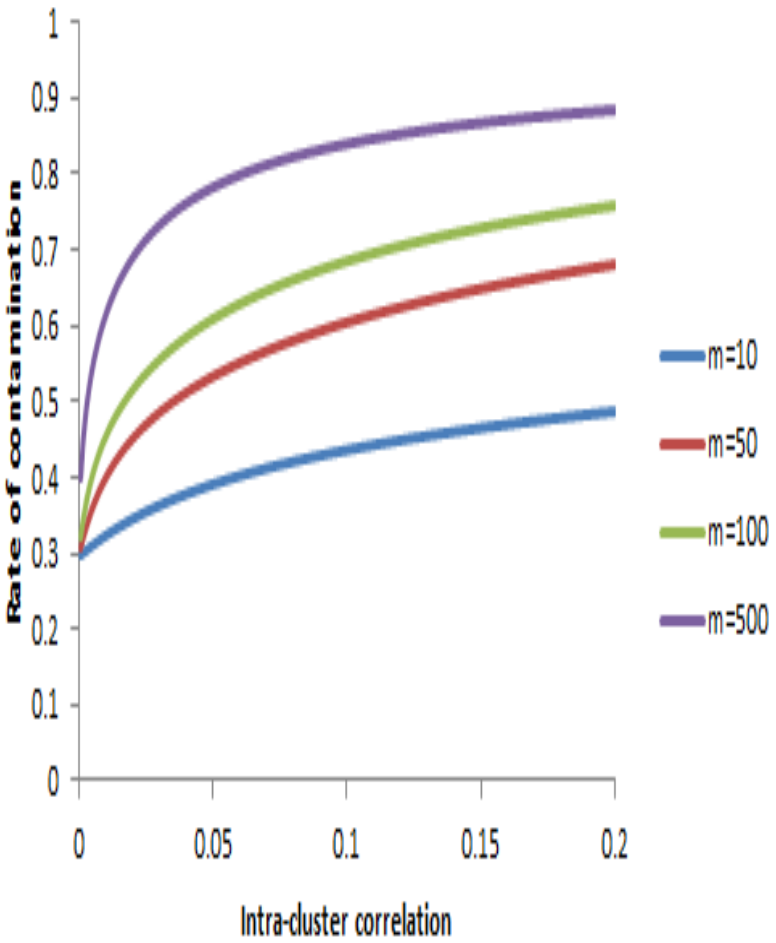
When $n_I* > n_{CRT-B}$. That is when:

$$2\sigma^2 \left[\frac{(z_{\alpha/2} + z_\beta)^2}{((1-w)\delta)^2} \right] > 2\sigma^2 \left[\frac{(z_{\alpha/2} + z_\beta)^2}{\delta^2} \right] 2[1 + (m-1)\rho](1-r^2) \quad (12)$$

That is, when:

$$(1-w)^{-2} > 2[1 + (m-1)\rho](1-r^2) \quad (13)$$

Critical values plot: CRT with baseline measures



Example (small cluster size)

For an ICC of 0.05, $CAC = 0.8$ and $m = 10$, up to 40% contamination can be tolerated before the sample sizes exceeds that of a CRT with a baseline.

Example (large cluster size)

For an ICC of 0.05, $CAC = 0.8$ and $m = 500$, up to 80% contamination can be tolerated before the sample sizes exceeds that of a CRT with a baseline.

Cluster randomised cross-over trial (CRXO)

Two period cluster cross-over trial

In a cluster cross-over trial clusters are allocated to one of two sequences. Clusters allocated to the first sequence are initially observed in the control condition and then switch to the intervention condition. And visa-versa.

The sample size under each treatment condition to detect δ, for (within-period) ICC ρ, cluster size m per period and where η is the CAC is:

$$n_{CRXO} = \underbrace{2\sigma^2 \left[\frac{(z_{\alpha/2} + z_{\beta})^2}{\delta^2}\right]}_{n_I} \underbrace{[1 + (m-1)\rho]}_{DE}(1-r) \qquad (14)$$

where $r = \frac{m\rho\eta}{1+(m-1)\rho}$ is the cluster mean correlation.

[1] Giraudeau B, Ravaud P, Donner A. Sample size calculation for cluster randomized cross-over trials. Stat Med. 2008 Nov 29;27(27):5578-85.

Comparing design effects: CRXO

How much contamination can be tolerated?

Determine the contamination rate at which an individually randomised trial with an attenuated treatment effect requires a larger sample size than a two period cluster cross-over design.

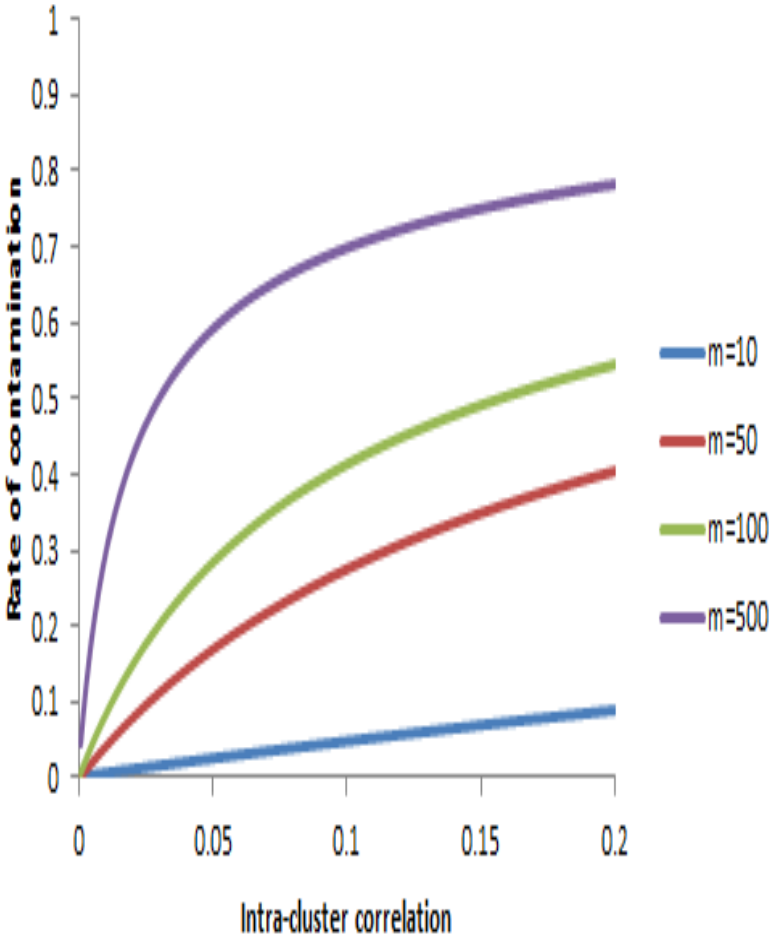
When $n_I* > n_{CRXO}$. That is when:

$$2\sigma^2 \left[\frac{(z_{\alpha/2} + z_\beta)^2}{((1-w)\delta)^2} \right] > 2\sigma^2 \left[\frac{(z_{\alpha/2} + z_\beta)^2}{\delta^2} \right] [1 + (m-1)\rho](1-r) \quad (15)$$

That is, when:

$$(1-w)^{-2} > [1 + (m-1)\rho](1-r) \quad (16)$$

Critical values plot: CRXO (CAC=0.8)



Example (small cluster size)

For an ICC of 0.05, $CAC = 0.8$ and $m = 10$, up to 2% contamination can be tolerated before the sample sizes exceeds that of a CRXO.

Example (large cluster size)

For an ICC of 0.05, $CAC = 0.8$ and $m = 500$, up to 60% contamination can be tolerated before the sample sizes exceeds that of a CRXO.

Compared to a stratified individually randomized trial

Stratified individually randomized trial

In a stratified individually randomized trial, individuals are allocated to either of two treatments at random, but such that within any given strata (i.e. centre in a multi-centre trial) there is a balance across treatment and control.

The required sample size per arm for a trial to detect a difference of δ, for intra-strata correlation (ISC) ρ_S is: n_{ST}, where:

$$n_{ST} = \underbrace{2\sigma^2 \left[\frac{(z_{\alpha/2} + z_\beta)^2}{\delta^2} \right]}_{n_I} [1 - \rho_S] \tag{17}$$

[1] Kahan BC, Morris TP. Analysis of multicentre trials with continuous outcomes: when and how should we account for centre effects? Stat Med. 2013 Mar 30;32(7):1136-49.

Comparing design effects: iRCT stratified vs CRT

How much contamination can be tolerated?

Determine the contamination rate at which a *stratified* individually randomised trial with an attenuated treatment effect requires a larger sample size than a CRT.

When $n_{ST} > n_{CRT}$. That is when:

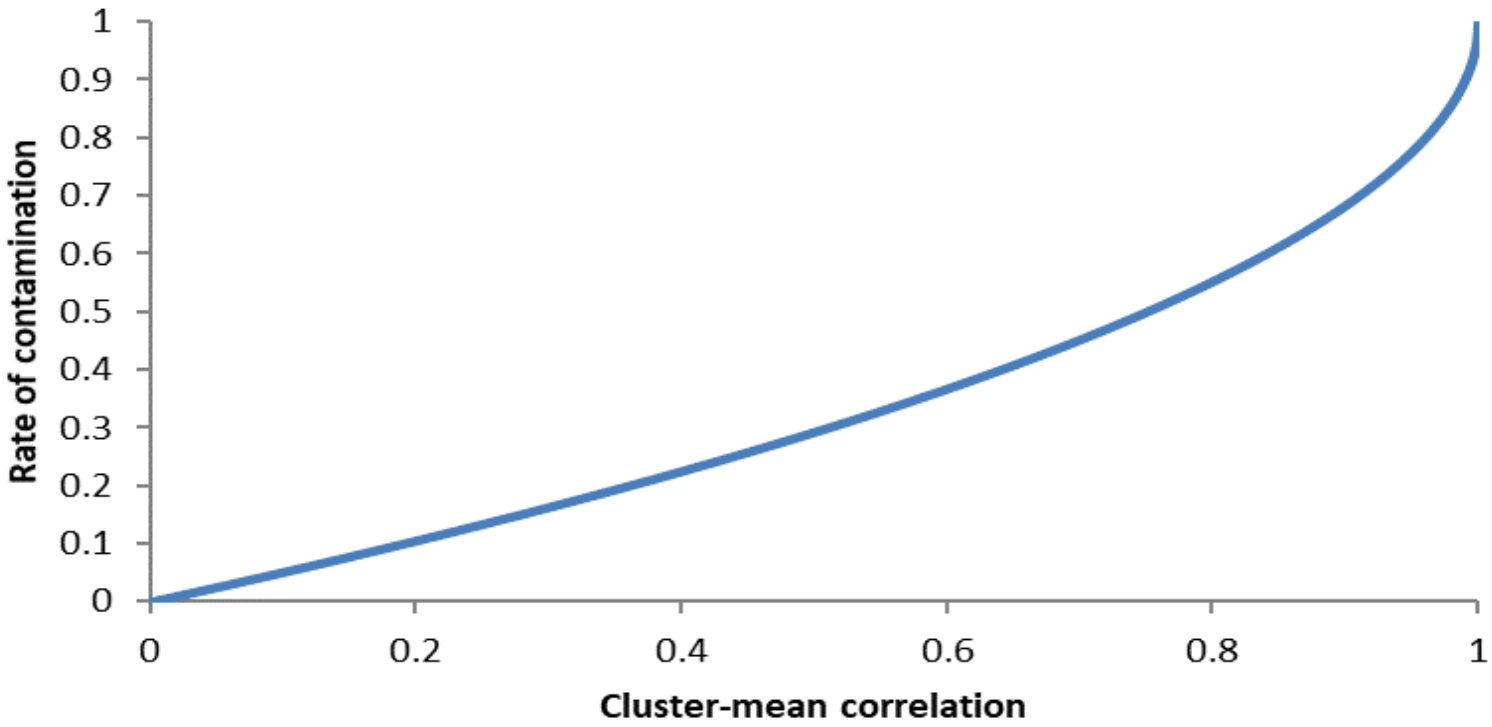
$$2\sigma^2 \left[\frac{(z_{\alpha/2} + z_{\beta})^2}{((1-w)\delta)^2}\right](1-\rho) > 2\sigma^2 \left[\frac{(z_{\alpha/2} + z_{\beta})^2}{\delta^2}\right][1+(m-1)\rho] \quad (18)$$

That is, when:

$$(1-w)^{-2}(1-\rho) > [1+(m-1)\rho]w > 1-\frac{[1+(m-1)\rho]^{-1/2}}{(1-\rho)^{-1/2}}w > 1-(1-r)^{1/2} \quad (19)$$

since $1-r = \frac{(1-\rho)}{[1+(m-1)\rho]}$ and assume $\rho = \rho_S$.

Critical values plot: *i*RCT stratified vs CRT



So it all depends on the cluster-mean correlation...

So, an individually randomised trial with stratified randomsation could tolerate larger amounts of contamination for larger cluster mean correlations ($r = \frac{m\rho\eta}{1+(m-1)\rho}$).

Critical values for different designs

Design	**Comparative**	**Critical value**
CRT	iRCT	$w > 1 - [1 + (m-1)\rho]^{-1/2}$
CRT-B	iRCT	$w > 1 - (2[1 + (m-1)\rho](1 - r^2))^{-1/2}$
CRXO	iRCT	$w > 1 - ([1 + (m-1)\rho](1 - r))^{-1/2}$
CRT	iRCT stratified	$w > 1 - (1 - r)^{1/2}$

So when rate of contamination is more than:

$$1 - [\textit{Design Effect For Clustered Design}]^{-1/2}$$

This is a critical value that, if exceeded, makes the individually randomized trial less "efficient" and less appealing.

Summary

- Rate of contamination that can be tolerated depends on within-cluster correlations (including cluster auto correlations) and cluster-size.
- Rate of contamination that can be tolerated increases with factors known to make the CRT less statistically efficient:
  - Large cluster sizes
  - Large ICC
  - Small CAC

  } i.e. high cluster mean correlation

Implications

- CRTs are generally at a greater risk of bias compared to individually randomised designs.
  - When an intervention is delivered at the level of the cluster, CRTs are impossible to avoid.
  - When the intervention is delivered directly to the individual, **individual randomisation is theoretically possible**.
- Individually randomized trials might still be the design of choice even in the presence of contamination:
  - Sample size can be increased to detect an attenuated effect.
  - Often this increase in sample size is much less than that needed to allow for clustered design.