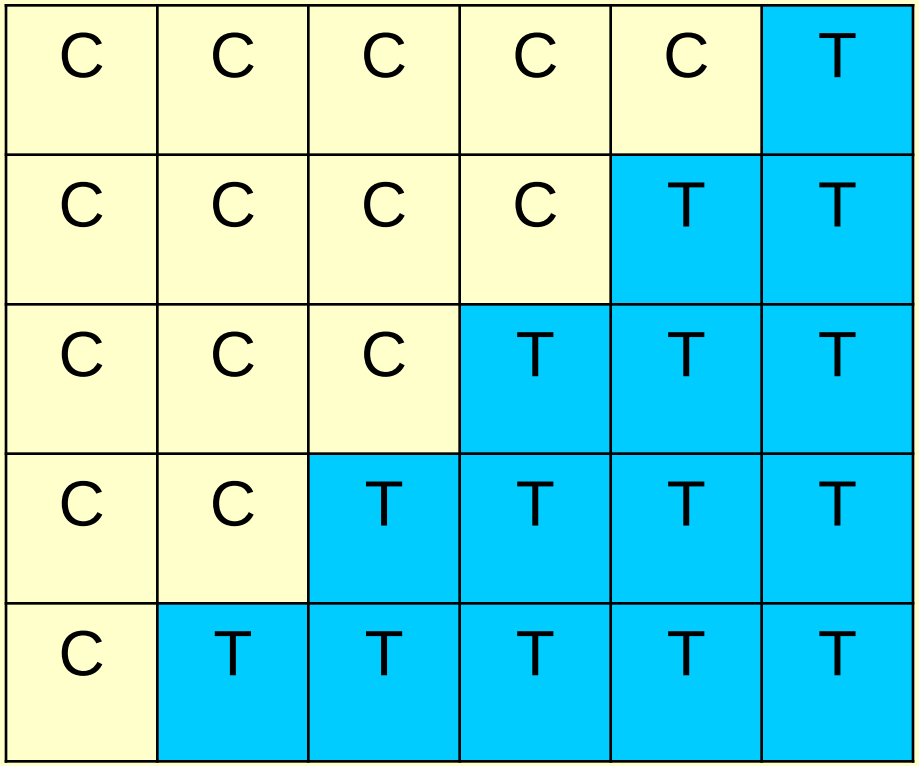
Stepped wedge randomised trials

Simon Thompson
University of Cambridge, UK

August 2011

What an	interesting	design !
	useful	
	challenging	
Mixture of	cross-over	trial
	cluster	
	longitudinal	

Randomized clusters (or individuals)



Time ⟶

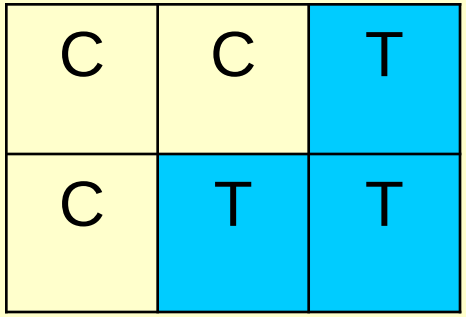
Vertical comparison

Horizontal comparison

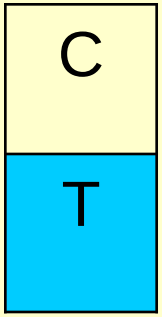
Baseline – all C

Final – all T

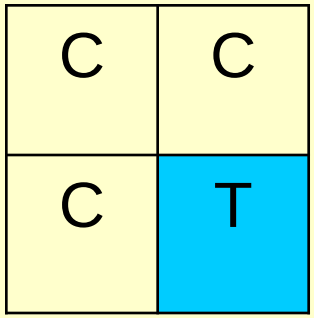
Varieties of cluster randomized trials



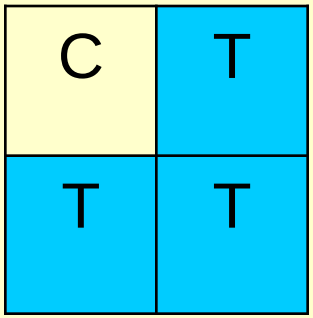
SW CRT



Parallel group CRT

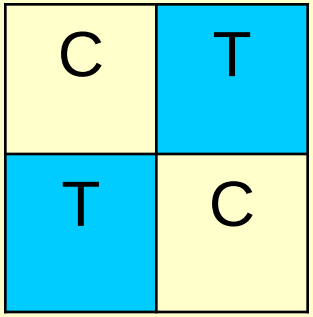


CRT with baseline

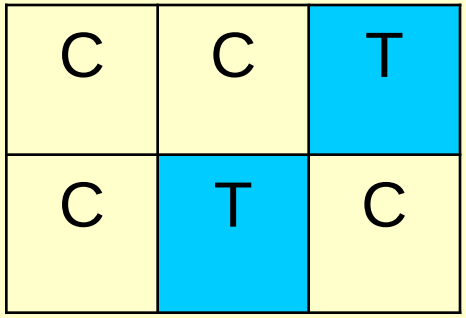


CRT with follow-up

Cross-over cluster randomized trials



Cross-over CRT



Cross-over CRT with baseline

Statistical advantages of SW design

1. Treatment effect is estimated both within and between clusters
2. Effect of between-cluster variability (ICC) is diminished
3. Can investigate the effect of duration of treatment

Mdege, Taylor et al, J Clin Epidemiol 2011

Review of 25 SW CRTs (studies and protocols)

Statistical analysis (plans) were often unclear / vague / inadequate

7 mentioned hierarchical / multilevel modelling
1 mentioned repeated measures
2 mentioned GEEs

9 “not stated”

Others: e.g. t-tests / chi-squared tests / linear or logistic regression, comparing pre- and post-intervention periods

Simple comparison of pre- and post-intervention periods:

Inadequate! Need to address:

- Clustering
- Repeated measures
- Confounding by time

Step-by-step assessment of intervention effect:

Helpful initial / exploratory analysis?

General basic analysis model

Y_{ijk} is continuous outcome for individual i in cluster j at step k

Model for SW cohort CRT (LMM)

$$Y_{ijk} = \alpha + u_j + \beta_k + v_{ij} + \theta\, T_{ijk} + e_{ijk}$$

where T_{ijk} = 1 if cluster j received treatment at step k
0 otherwise

α, β_k and θ are fixed effects

$u_j \sim N(0,\sigma_u^2)$ are random cluster effects

$v_{ij} \sim N(0,\sigma_v^2)$ are random individual effects

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

Adaptations

(1) Binary outcomes: $Y_{ijk} = 0 / 1$ (GLMM)

$$Y_{ijk} \sim \text{Bin}(1, \pi_{ijk})$$

$$\text{logit}(\pi_{ijk}) = \alpha + u_j + \beta_k + v_{ij} + \theta\, T_{ijk}$$

(2) No clusters j (SW individually randomized trial)

$$Y_{ik} = \alpha + \beta_k + v_i + \theta\, T_{ik} + e_{ik}$$

(3) SW repeated cross-sectional CRT

$$Y_{ijk} = \alpha + u_j + \beta_k + \theta\, T_{jk} + e_{ijk}$$

Extensions

Duration m

0	0	0	0	0	1
0	0	0	0	1	2
0	0	0	1	2	3
0	0	1	2	3	4
0	1	2	3	4	5

(1) Treatment duration:

(e.g. delayed treatment effect)

Replace θ by θ_m

(2) Cluster variability σ_u^2 depends on T=0/1

Individual level variability σ_v^2 or σ_e^2 depends on T=0/1

(3) Incorporating:

random cluster x time interaction

random cluster x treatment interaction

Need a statistical analysis plan / explicit statistical model !

Analysis strategies

Hierarchical modelling is a natural approach

Cluster-level models are an option, especially for equally sized clusters

GEEs not good for multiple levels of clustering, or small numbers of clusters?

Advantages of a Bayesian / MCMC approach:

propagates full uncertainty of variance components

can use informative priors for ICC

flexible framework for modelling extensions

Other problems

Missing data / attrition

Survival outcomes: differential attrition

Non-compliance of individuals / clusters

Contamination

Lack of blinding

Use of cluster / individual level covariates

Adjusting for baseline

Model checking

Conclusions

Write down the statistical analysis plan

Give principal statistical analysis model explicitly

Mention proposed extensions / sensitivity analyses

Base design / power calculation on principal statistical analysis model (analytically / by approximation / by simulation)