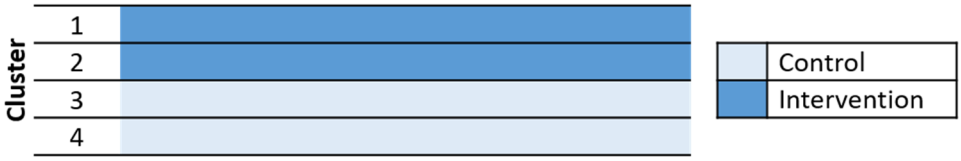
A Tour of Pragmatic Study Designs: Cluster Randomized Trial

Miriam Dickinson, PhD; Allison Kempe, MD, MPH

A **Cluster Randomized Trial (CRT)** is a trial in which clusters (e.g., hospitals, regions) rather than individuals are randomized to different intervention groups. A key implication of cluster randomization is that the responses of multiple individuals in the same cluster are usually positively correlated. Due to this positive intracluster correlation, advanced statistical methods (such as mixed models) must be considered for analysis.

Parallel Cluster Randomized Trial. Clusters are randomized to either the intervention or control arm at the start of the trial and remain in that arm for the remainder of the study.



The **advantages** of a CRT are that it is a simple design that is easy to implement. It is often considered when randomization at the individual level is not possible. Cluster randomization protects against contamination across intervention groups when patients are managed within the same setting or by the same provider. The **challenges** of a CRT are that it requires a large number of clusters to detect small effect sizes with adequate power. In a trial with a true control group, not all clusters will receive the intervention during the study.

Correlation in CRTs. An important implication in CRTs is that patients within a single cluster are often more likely to respond similarly due to physical, geographic, and social commonalities, and thus cannot be considered to contribute independent observations. This lack of independence results in a loss of statistical power compared to trials randomized at the individual level. To quantify how strongly patients in the same cluster resemble each other, the statistical measure **intracluster correlation coefficient (ICC)** is used. To achieve equivalent power to a patient randomized trial, standard sample size calculations must be inflated by a factor of

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

where m is the average cluster size, and ρ is an estimate of the ICC. This is referred to as a **design effect.**

Estimating ICC. ICC takes a value between 0 and 1, where an ICC closer to 1 indicates that there is high similarity between responses from individuals in the same cluster. ICCs for disease outcomes are generally less than 0.05. ICC can be estimated from other trials with similar populations and endpoints.

The **analysis** of data collected from a CRT must also account for clustering. Analyses at the cluster-level (i.e., using summary measures for each cluster) are generally not statistically efficient. Patient-level analysis can account for clustering using **generalized linear mixed models (GLMM)** and **generalized estimating equations (GEE)**. These modeling techniques also allow for the adjustment of both cluster-level and patient-level covariates.

ACCORDS
ADULT AND CHILD CONSORTIUM FOR HEALTH OUTCOMES RESEARCH AND DELIVERY SCIENCE
UNIVERSITY OF COLORADO | CHILDREN'S HOSPITAL COLORADO

COPRH Colorado Pragmatic Research in Health Conference 2020

A Tour of Pragmatic Study Designs: Cluster Randomized Trial

Miriam Dickinson, PhD; Allison Kempe, MD, MPH

Notes:

ACCORDS
ADULT AND CHILD CONSORTIUM FOR HEALTH OUTCOMES RESEARCH AND DELIVERY SCIENCE
UNIVERSITY OF COLORADO | CHILDREN'S HOSPITAL COLORADO

CoPRH Colorado Pragmatic Research in Health Conference 2020