Spatial Cluster Randomized Trials - Sampling Design with Spillover Effects & Spatial Dependence

Andrew Walther[1], Tonya Van Deinse[2], and Feng-Chang Lin[1]

[1]The University of North Carolina at Chapel Hill, Department of Biostatistics

[2]The University of North Carolina at Chapel Hill, School of Social Work

*Corresponding author: flin@bios.unc.edu

January 28, 2026

Keywords: spatial autoregressive, simple random sampling, block stratified sampling, treatment assignment, clinical trials

Abstract

Background: Investigators designing cluster randomized trials desire insights directing treatment assignment methodology for studies involving spillover effects and spatial dependence.

Methods & Simulation: Treatment assignment strategies including simple random sampling (SRS) and block stratified sampling (BSS) are defined and spatial autoregressive modeling is applied with consideration for spillover effects and spatial dependence for estimation of intervention effects. A simulation study is carried out comparing SRS and BSS sampling methods on spatial grids of varying sizes. A range of spillover effects and levels of spatial dependence were considered for estimation of the intervention effect via a spatial autoregressive (SAR) model.

Results: Findings of an extensive simulation study comparing simple random sampling and block stratification methods indicate that randomly selected treatment assignments result in best case reduced Mean Squared Error (MSE) when estimating intervention effects, but block stratified treatment assignments lead to minimal variation in MSE among a series of treatment combinations, indicating that a block stratified treatment arrangement won't achieve the minimal level of estimation error, but it remains robust across a range of selected parameters. Even though SRS is consistent

and unbiased with reduced MSE, we consider variation among all possible treatment combinations to ensure a robust result.

Conclusion: The relationship between spillover effects and mean squared errors (MSE) of intervention effect estimation is apparent. The MSE for the intervention effect, which is the average MSE over each of N simulation iterations, is minimized for some combinations of random sampling treatment assignment, but block stratified assignment minimizes variance among combinations of possible treatment arrangements. In short, the SRS technique may achieve minimum average MSE in some cases, but BSS achieves respectable average estimation error with minimal variation between treatment combinations.

1 Introduction

1.1 Background & Motivation

Randomized controlled trials (RCTs) represent the gold standard for establishing causal inference when assessing the efficacy and safety of new interventions[1][2]. However, in many real-world scenarios, particularly in public health, education, and the social sciences, randomizing individuals is not feasible due to logistical constraints, ethical considerations, or the inherent nature of the intervention itself, which may be administered at a community level. In such cases, the cluster randomized trial (CRT) emerges as the preferred and most rigorous design, where intact social units or groups such as villages, schools, or clinics are randomized to intervention or control arms[3][4].

A key advantage of the CRT design is its ability to evaluate interventions where there is high risk of contamination between treated and untreated individuals within the same cluster[3][5]. Furthermore, CRTs are uniquely positioned to allow for the estimation of not only the direct effects of an intervention on those who receive it but also the indirect spillover effects. These are the additional impact of an intervention on individuals who did not directly receive it, typically including those in control clusters. Measuring these spillover effects is often of primary scientific and policy interest, as it provides a comprehensive assessment of an intervention's total public health impact[3][6]. When trial clusters are defined by geographic areas, as is frequently the case, the design must contend with inherent spatial complexities that can challenge the core statistical assumption of cluster independence. Two primary and interrelated spatial phenomena can profoundly influence trial outcomes and compromise the validity of the results:

spillover effects and spatial dependence[3][2].

Spillover occurs when an intervention's effects in one cluster extend or "spill over", into neighboring clusters, creating significant analytical challenges[7][8]. This is especially common in trials of infectious disease interventions. A study on the use of permethrin-impregnated bed nets by Binka et al. (1998) provided early, compelling evidence of this phenomenon, demonstrating that child mortality was significantly lower in control villages situated near intervention villages, presumably due to a reduction in the local mosquito population[9][10][11]. Similarly, the work by Miguel and Kremer (2004) on a school-based de-worming program in Kenya found significant positive externalities, showing that the health and school attendance of untreated children improved if they lived near schools that were part of the intervention[12]. The critical implication is that failing to account for such positive spillover can lead to a substantial underestimation of the intervention's true effect, as the measured outcomes in the control group are artificially improved by their proximity to the intervention arm. Compounding this issue is spatial dependence, a fundamental concept supported by Tobler's First Law of Geography that asserts nearby observations tend to be more related than distant ones[8][13]. In the context of a CRT, this principle suggests that outcomes in nearby clusters may be correlated due to unmeasured, spatially-patterned environmental, social, or demographic factors. The presence of this spatial autocorrelation violates the assumption that clusters are independent units, which can lead to biased standard errors and an increased Type I error rate undermining statistical inference[14][15][13].

In response to these challenges, a body of literature has emerged focusing on post-hoc analytical adjustments to account for spatial effects after the data have been collected. A systematic review by Jarvis et al. (2017) found that these existing analytical approaches generally fall into two distinct categories: the spatial variable approach and spatial modeling[8]. The spatial variable approach is the more common of the two. It involves creating a new covariate based on distance or density metrics. Examples include measuring the straight-line distance from control participants to the nearest intervention household (Hawley et al., 2003 & Gimnig et al., 2003)[11][10], calculating outcome rates across different distance bands (Binka et al., 1998)[9], or quantifying the density of treated individuals within a predefined radius (Miguel & Kremer, 2004 & Lenhart et al., 2008)[12][16]. More recently, Chao et al. (2015) developed a "potential exposure" metric to quantify the spatially heterogeneous risk surrounding an individual[17]. In contrast, the spatial modeling approach is more statistically sophisticated, directly

incorporating the geographic structure of the data into the random effects component of a statistical method. The method employed by Alexander et al. (2003)[18] considers a distance-decay function in the covariance matrix and an alternative by Silcocks and Kendrick (2010)[19] using spatial weights matrices treats spatial correlation as an underlying unobserved process. Additionally, Watson & Smith (2025) consider an alternative framework from cluster-focused intervention application and instead randomly assign intervention areas and measure "treatment" as a continuous dose which is a function of the observation's distance to the nearest intervention site[20]. Estimation of a dose-response function (DRF) describes how the treatment effect varies over distance from the intervention source rather than directly estimating the intervention or spillover effect in a cluster-based trial.

Despite the growing sophistication of these analytical tools, they are fundamentally post-hoc corrective solutions to a problem that could (and should) be addressed proactively at the design stage. The existing literature reveals a noticeable lack of research into how the initial sampling and allocation of treatments across a geographic landscape can be optimized to manage these known spatial complexities. The review by Jarvis et al. (2017) highlights this gap, noting that in practice, the impact of location on trial results is definitely an afterthought and oftentimes entirely ignored[8]. In fact, trial designs explicitly intended to mitigate spatial effects, such as creating well-separated clusters or buffer zones, were so uncommon that studies employing them were excluded from their primary analysis, underscoring how peripheral these considerations have been to standard practice[8].

This oversight has significant consequences because known studies that compare spatial analytical approaches consistently find that accounting for spatial effects is important. Silcocks and Kendrick (2010) demonstrated that spatial models not only fit primary care trial data better than standard CRT models, but also that their inclusion altered both the intervention effect point estimate and its precision[19]. Similarly, Chao et al. (2015) found that adjusting for their spatial exposure metric lowered the estimated effect of a typhoid vaccine[17]. The logical conclusion is that standard trial designs that rely on the simple random allocation of clusters, may be systematically producing biased or inefficient estimates without investigators being aware. This leads to the central question motivating the work of our study. That is, "How does the choice of an *a priori* treatment assignment sampling strategy impact the estimation of intervention effects in the presence of both spatial spillover effects and spatial dependence?" While we have methods to analyze the problem after a study is completed, there is currently a lack of evidence-

based guidance for investigators on how to design a cluster randomized trial to mitigate these issues from the outset of study planning.

1.2 Aims & Contributions

This study aims to address this critical gap by moving beyond post-hoc analysis to investigate the important role of treatment sampling design in spatial CRTs. Through a comprehensive simulation study, we systematically compare the performance of two contrasting approaches for assigning study clusters to intervention and control arms: simple random sampling (SRS), representing the standard practice for treatment assignment, and a block stratified sampling (BSS) method. The BSS strategy is intentionally designed to create spatial separation between clusters of the same treatment assignment by isolating them to potentially reduce the confounding influence of spillover and spatial dependence.

Our simulations explore how the estimation error of the intervention effect, quantified by Mean Squared Error which is composed of bias and variance, varies across a range of realistic conditions including different magnitudes of spillover effects, the presence or absence of underlying spatial dependence, and variation of restrictions on the geographic distance over which spillover can occur. By identifying the systematic patterns that emerge, this novel research provides practical, evidence-based guidance for investigators designing future spatial CRTs. The findings will equip researchers with the knowledge to select sampling strategies that align with their specific study goals, ultimately leading to more robust, efficient, and valid estimates of intervention effects in complex spatial settings.

The remainder of this article is organized as follows: Section 2 outlines the theoretical framework of the study, including the data generation process, treatment sampling procedures, and estimation methods employed. Section 3 details the procedure and results of the simulation study undertaken to evaluate the sampling methods considered. Section 4 considers an application of these findings to make a recommendation for other investigators. Finally, Section 5 discusses the implications of our findings, potential drawbacks to the work performed, and potential areas for continued future research.

2 Design of Spatial CRTs

2.1 Cluster Randomized Trials

The randomized controlled trial (RCT) serves as the gold standard design to evaluate intervention efficacy[21]. Its core principle, the random allocation of individual participants, is expected to balance both measured and unmeasured confounding variables across treatment arms[22]. However, individual randomization is often impractical, or an intervention is delivered at a group level, necessitating an alternative approach[4][3]. In these scenarios, the cluster randomized trial (CRT) is employed, where intact social units or groups like households, clinics, or geographic areas are randomized to intervention arms[4][23]. A key advantage of this design is the prevention of treatment contamination, where control group members are inadvertently exposed to the intervention, which would otherwise bias the estimated treatment effect toward the null[3]. Statistical considerations in a CRT are the measured response from an intervention and the inherent correlation among outcomes of individuals within the same cluster, a dependency quantified by the intraclass correlation coefficient (ICC)[4][24]. The ICC measures how similar the outcomes of individuals within a cluster are likely to be, relative to those of other clusters. This intra-cluster correlation violates the standard assumption of independence between observations. The failure to account for this feature in the analysis leads to underestimated standard errors and an inflated probability of a Type I error[25][26]. Therefore, the valid design and analysis of CRTs require specialized statistical methods that properly account for this clustering structure, with comprehensive frameworks provided by Donner & Klar (2000)[4] and Hayes & Moulton (2017)[3].

2.2 Neighbor Definitions

2.2.1 Contiguity-Based

Contiguity-based neighbors[27] are constructed with the assumption that neighbors of a given area (individual cells in the case of this study) are other clusters (or cells) that share a common boundary. This boundary may be an edge and/or vertex. The **spdep** R package[28][29][30][31] includes functions for working with spatial neighbors and spatial dependence structures. In this study, we consider Rook Neighbors (shown in Figure 1 (a)), which fulfill a contiguity condition where a pair of clusters share a common boundary made up by an edge. Queen neighbors (shown in Figure 1 (b)) are another type of

contiguity-based neighbors that fulfill a contiguity condition where a pair of clusters share a common boundary of either an edge and/or a vertex[27]. Clusters may also have Bishop neighbors where a common boundary at the vertex (corner) of the spatial region is shared. Clusters may also have higher-order neighbors where they do not directly share an edge or vertex, but rather have a common boundary with a cluster that is a neighbor of another cluster. In other words, these are "neighbors of neighbors" of the reference cluster.

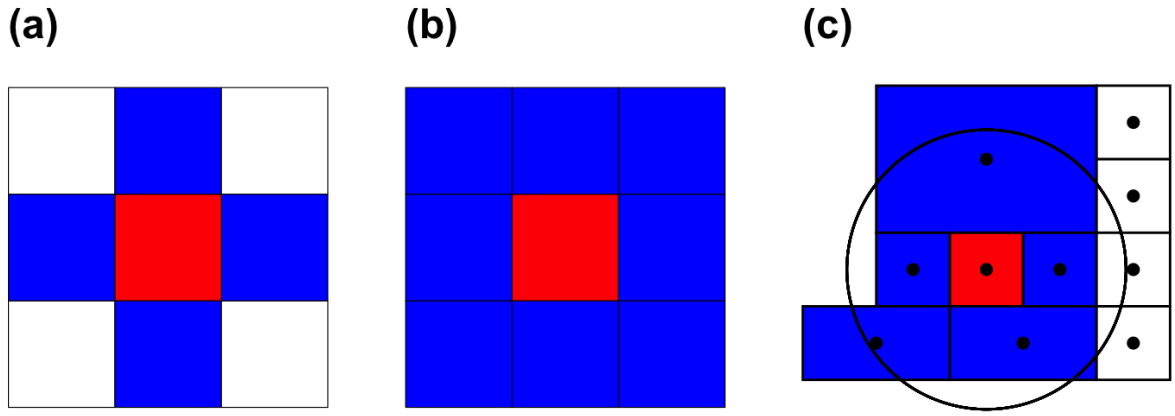

Figure 1: (a) Rook neighbors (blue) share an edge with a reference spatial region (red), (b) Queen neighbors (blue) share either an edge or vertex (or both) with a reference spatial region (red), and (c) Distance-based neighbors (blue) have centroids within a fixed radius of the reference cluster (red). Clusters with centroids beyond this distance are not considered as neighbors (white)[32]

2.2.2 Distance-Based

Distance-based neighbors[27] are constructed with the assumption that neighbors of a given area (individual cells in the case of this study) are other areas (or cells) that lie within a specified distance of another cell or observation. This distance is commonly defined by the radius of a circle around the center of a reference spatial region. The region specified in Figure 1 (c) identifies distance-based neighbors of the reference cell (dark) as neighboring cells where the centroid of the cells is within the circumscribed boundary around the reference cell. In addition to cells, distance-based neighborhoods can also be constructed based on individual points that are within a specified distance of another point.

2.3 Intervention & Spillover Effects

Two primary effect types are considered to understand how treatments affect subjects in clusters where they are applied and subjects in the surrounding areas. The intervention effect and spillover effect are two parameters that are estimated in this study. Spatial cluster randomized trials have subjects or regions that are assigned to receive an intervention and others assigned as controls that do not receive an intervention. Observation of the resulting outcomes allows for estimation of the direct intervention effect and indirect spillover effect on the study region.

2.3.1 Intervention Effect

The intervention effect (β) represents the estimated effect of the direct intervention applied in a study. This is a measure of the study outcome that is directly attributed to the intervention applied in the study. Estimation of this intervention effect is possible due to the presence of clusters assigned as a control for comparison. The relevant intervention could be a vaccine to prevent infectious disease, a training program for government employees, an educational program for students, or many others.

2.3.2 Spillover Effect Methods

The spillover effect (ψ) represents an indirect effect on clusters where an intervention was not directly applied[33][34]. There are many to model a spillover effect as it is observed due to spatial proximity to subjects or regions where an intervention is applied. In this study, only binary spillover effects were considered. A binary spillover effect is a fixed level of a spillover effect that is applied if a subject or region is deemed to be a neighbor of a subject or region where the intervention is applied.

2.3.3 Spillover Effect Types

In this study, there are two specific ways in which spillover of the applied intervention is allowed to propagate to other cells or subjects. In both cases, spillover effects originate from areas where the original intervention treatment is applied. The first case allows for spillover to propagate from cells receiving the intervention into only cells assigned as control cells. In this case, the effect of the intervention is not permitted to spillover into other intervention cells. This suggests a case where the intervention effect applied is the maximum observable effect and any potential spillover cannot increase the observed effect of a particular intervention. Another method of spillover application is considered where spillover

effects are permitted to be applied to both control AND intervention cells that are neighbors of a cell that receives an intervention. This application suggests that an effect greater than that of the original intervention applied can potentially be observed in a cell that received an intervention and also had spillover from another intervention cell. It is also to be expected that any change in the estimated effects would not be observed in spatial settings where the treatment allocation does not satisfy block stratification conditions (where no intervention cells are rook neighbors with other intervention cells). Figure 2 provides a comparative representation of how these two methods permit the application of spillover effects on clusters based on their treatment assignment. In all scenarios of this simulation study, the application of any spillover effect is simply a binary indication of whether a cluster receives a spillover effect or not. Even if spillover exposure could potentially be directed from multiple neighboring clusters, the modeling in this study does not consider any additive spillover effects.

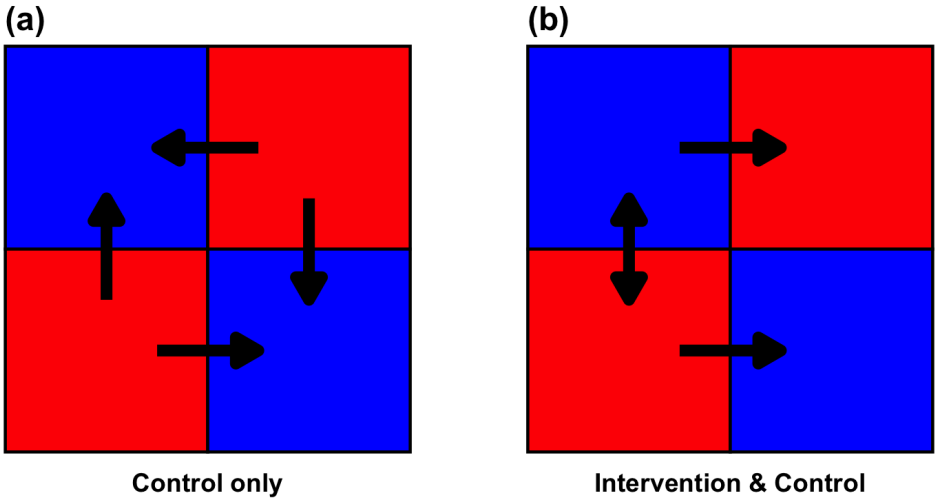

Figure 2: (a) Spillover application from intervention cells (red) to only control cells (blue) & (b) spillover application to intervention (red) and control cells (blue).

2.4 Simple Random vs. Block Stratified Sampling

In this study, two different treatments are considered: intervention & control. Assignment of these treatments is considered under two sampling methods: Simple Random Sampling (SRS) & Block Stratified Sampling (BSS). Random sampling treatment assignment considers all possible combinations of intervention & control arrangements on a specified spatial grid and provides equal probability of assignment for any given combination of intervention and control. For example, a 2 row by 3 column spatial grid has

6 total cells. Consideration of a 1:1 ratio between intervention and control treatments provides that three cells are assigned as intervention and three cells are assigned as control. Therefore, there are ${}_6C_3 = 20$ different combinations of treatments for this specific spatial region.

Block stratified sampling ensures a balanced number of intervention and control cells are selected, with the added condition that no intervention or control cells share a common edge boundary. An exception to this rule occurs in spatial setups where both dimensions are odd. In this case, the count of intervention and control cells must be out of balance and we assert that the control class will always be in the majority. In these instances, we relax the common edge boundary restriction for control cells and only require that intervention cells not share a common edge boundary in order to satisfy the block stratification condition.

We now consider a series of spatial grid setups to address efficient estimation of the intervention effect. These spatial grids are have dimensions denoted as (row $\times$ column) and we seek to apply a balanced distribution between intervention and control cells where possible. When this isn't possible, we defer to the control group being the majority class as mentioned previously.

2.4.1 2x4 Grid

One representative spatial setting considered is grid of 8 individual cells laid out in an arrangement of 2 rows and 4 columns. For this grid setup, a balanced set of intervention and control cells can be assigned where 4 cells are intervention regions and 4 cells are control regions. For a grid of 8 total cells and 4 cells of each treatment type, there are ${}_8C_4 = 70$ unique arrangements of treatment and control assignments. Figure 3 (a) displays one of these combinations that satisfies the block stratified condition where intervention and control cells do not have any rook neighbors of the same treatment type. For this spatial setting, there are 2 (out of 70) combinations that satisfy the block stratification condition. Figure 3 (b) displays an example of the 2x4 grid spatial setting with a randomly sampled arrangement of intervention and control cells. Note that both treatment types have cells that are rook neighbors with the same treatment. These spatial grids are constructed with Simple Features for R package[35][30].

2.4.2 3x3 Grid

The 3x3 symmetric grid setting of 9 cells allows for ${}_9C_4 =_9 C_5 = 126$ unique combinations of intervention and control cells where 5 cells are allocated as control and 4 cells are treatment regions. The majority class

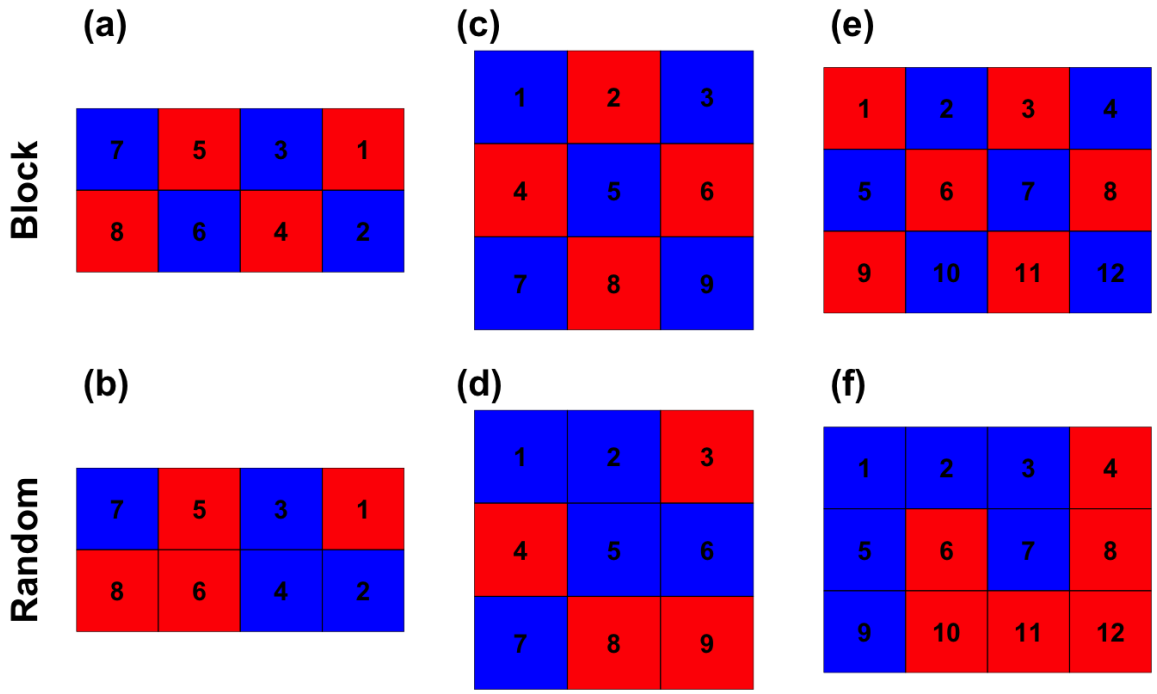

Figure 3: 2x4 (left), 3x3 (middle), and 3x4 (right) spatial grids of intervention (red) and control (blue) clusters assigned via Block Stratified Sampling (top row) and Random Sampling (bottom row)

of 5 cells is the control group for the sake of hypothetical resource management. Figure 3 (c) illustrates one combination of treatment and control cells that completely satisfies the block stratification conditions. Additional cluster combinations satisfy a relaxed condition of block stratification where control cells are allowed to have other control cells as rook neighbors. This is possible in spatial grid settings where both dimensions (row, column) are odd in length. There is one combination for the 3x3 grid setting that satisfies the strongest block stratification condition and 5 more combinations that satisfy the relaxed block stratification condition for a total of 6. An example of random sampling treatment assignment that doesn't satisfy block stratification conditions is provided in Figure 3 (d).

2.4.3 3x4 Grid

The final spatial setting considered was a 3 row by 4 column (3x4) grid of 12 total cells. This grid, with a balanced allocation of 6 intervention cells and 6 control cells allows for ${}_{12}C_6 = 924$ unique arrangements of the two treatments in the spatial region. For this particular grid size, 2 of the 924 treatment combinations satisfy the block stratification condition and Figure 3 (e) presents one of these, with a "checkerboard" design. The remaining treatment combinations are achieved via random sampling of a balanced number

of intervention and control cells. Figure 3 (f) illustrates another randomly sampled arrangement of the treatments onto the spatial area where control and intervention cells each have rook neighbors of the same treatment type.

2.5 Spatial Autoregressive (SAR) Model

Outcomes in the presence of spatial dependence and spillover effects can be considered via modeling with a spatial autoregressive

model[36][27][37]. A spatial autoregressive model has the form:

$$y_{ik} = \rho \sum_{j} w_{ij} y_{jk'} + \mathbf{x}_i \beta + \varepsilon_i \tag{1}$$

where k and k' are different cluster indices for each individual, ρ is the spatial correlation (autoregressive) coefficient, and ε_i represents an i.i.d. error term. This can also be expressed in matrix notation as:

$$\mathbf{y} = \rho \mathbf{W} \mathbf{y} + \mathbf{X} \beta + \varepsilon \tag{2}$$

It follows that the general form of the response for this model for the purpose of this study is

$$y_{ik} = \alpha + \beta x_{ik} + \psi z_{ik} + \rho \mathbf{W} \mathbf{y} + \epsilon_i \tag{3}$$

where y_i is the observed response for subject i, α is the baseline effect, β is the intervention effect, x_i is an indicator of whether subject i receives the intervention, $\psi(d_i)$ is the spillover effect expressed as a function of distance from the nearest intervention source for subject i, z_i is an indicator of whether subject i can experience any spillover effect, and $\rho\mathbf{W}\mathbf{y}$ captures spatial dependence from neighboring observations and their effects in the response variable. The components of $\rho\mathbf{W}\mathbf{y}$ are ρ, a fixed spatial correlation parameter, $\mathbf{W}$, the spatial weights matrix indicating the strength of relationship between neighboring observations, and the response vector $\mathbf{y}$. Finally, ε_i is a random error term for each subject.

2.6 Spatial Dependence

Spatial dependence, used interchangeably with spatial autocorrelation, refers to the property of spatial data where observations close to each other tend not be independent of each other. Rather, observations that are close to each other tend to be dependent on each other and influence values of surrounding

data. Hence, there is an implied systematic correlation between data at neighboring locations. Spatial correlation can be positive, where similar values tend to be clustered close together, or negative, where dissimilar values are close together. Appropriate consideration of spatial dependence is necessary for performing valid statistical inference as well as for accurate estimation of effects that are modeled. A pair of common methods for assessing spatial autocorrelation are Moran's I [38][39] and Geary's C [40][36]. Moran's I standardizes the spatial auto-covariance by the variance of the data and Geary's C utilizes the sum of the squared differences between pairs of data as a measure of co-variation. These test statistics are reliant on the spatial weights matrix.

The spatial weights matrix $\mathbf{W}$ is a component of the SAR model from Section 2.5. It is composed elements w_{ij} which indicate the strength or presence of a spatial connection between observations i and j. Therefore the spatial dependence term of the SAR model is represented as

$$\rho \mathbf{W}\mathbf{y} = \rho \sum_{j \in \mathcal{N}(i)} w_{ij} y_j \tag{4}$$

where ρ is the spatial autocorrelation parameter; an indicator of how strongly neighboring values affect the focal observation, $\mathbf{W}$ is the spatial weights matrix between subjects i and j with matrix elements w_{ij} that are either indicator-based ($\{0, 1\}$) or distance-based $\left(\frac{1}{d_{ij}}\right)$ where d_{ij} is the distance between subjects i and j, and $\mathcal{N}(i)$ is the set of all neighbors of observation i. In this study, this set consists of all subjects that are a member of a cluster that is a rook neighbor of the cluster of subject i. For the spatial weights matrix, The indicator-based weights are

$$w_{ij} = \begin{cases} 1, \text{if points } i \text{ and } j \text{ are neighbors} \\ \\ 0, \text{otherwise} \end{cases} \tag{5}$$

and are normalized via

$$W_{ij} = \frac{w_{ij}}{\sum_j w_{ij}} \tag{6}$$

and the distance-based weights are

$$w_{ij} = \begin{cases} \frac{1}{d_{ij}}, \text{if } 0 < d_{ij} < d_{\max} \\ \\ 0, \text{otherwise} \end{cases} \tag{7}$$

which are normalized via

$$W_{ij} = \frac{\frac{1}{d_{ij}}}{\sum_j \frac{1}{d_{ij}}}. \tag{8}$$

2.7 Estimation for SAR Model

Maximum likelihood estimation is utilized to gather parameter estimates from the spatial lag model from Section 2.5[41][42][43]. When considering spatial dependence, the responses are

$$y_i = \rho \sum_{j \neq i} w_{ij} y_j + \alpha + \beta x_i + \psi z_i + \epsilon_i \tag{9}$$

where the residuals are

$$\varepsilon_i = y_i - \rho \sum_{j \neq i} w_{ij} y_j - (\alpha + \beta x_i + \psi z_i)\,. \tag{10}$$

Then, assuming that $\varepsilon_i \sim N(0, \sigma^2)$, the likelihood function is

$$L\left(\alpha, \beta, \psi, \rho, \sigma^2\right) = \frac{1}{\sqrt{(2\pi\sigma^2)^n \left|\mathbf{I} - \rho\mathbf{W}\right|}} \exp\left(-\frac{1}{2\sigma^2}\epsilon^\top \epsilon\right) \tag{11}$$

where $|\mathbf{I} - \rho\mathbf{W}|$ is the determinant of $(\mathbf{I} - \rho\mathbf{W})$. The log-likelihood is

$$\log L\left(\alpha, \beta, \psi, \rho, \sigma^2\right) = -\frac{n}{2}\log\left(2\pi\sigma^2\right) + \log\left|\mathbf{I} - \rho\mathbf{W}\right| - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i - \rho\sum_{j \neq i} w_{ij} y_j - (\alpha + \beta x_i + \psi z_i)\right)^2 \tag{12}$$

which is maximized with respect to $\alpha, \beta, \psi, \rho,$ and σ^2. In applied simulation tasks, estimates can be obtained via the 'lagsarlm()' function from the R spdep library[28].

3 Simulation Study

A simulation study was carried out to consider the contrast of block stratified and random treatment assignment sampling methods with consideration for spillover effects, spatial dependence, and restrictions on spillover application between intervention and control regions. The results of this study are detailed in the remainder of this section.

3.1 Simulation Scenarios & Response

We carry out a simulation study to explore how estimation error of the intervention effect in a Spatial CRT varies across a range of spillover effect sizes, the presence or absence of spatial dependence, varying restrictions on clusters where spillover is permitted to occur, and the method of sampling for treatment assignment. Each simulation scenario is composed of a set of these parameters. The baseline effect, α,

is always assigned as 0.2. The intervention effect, β, is always assigned as 1.0. The spillover effect, ψ, has possible values 0.5, 0.6, 0.7, and 0.8. The spatial dependence parameter, ρ is set as either 0.00 or 0.01. The sampling design for each scenario is either Block Stratified Sampling (BSS) or Simple Random Sampling (SRS). Finally, the spillover application type is (1) Control Clusters Only or (2) Control & Intervention Clusters. Each simulation setting is replicated for each of the spatial grid dimensions discussed previously (2x4, 3x3, and 3x4).

Given that the form of the spatial lag model in Section 2.5 contains the response on both sides of the expression, a separate closed-form expression for data generation must be considered[27][44]. This data generation expression is

$$\mathbf{y} = (\mathbf{I} - \rho\mathbf{W})^{-1} (\boldsymbol{\alpha} + \mathbf{X}\boldsymbol{\beta} + \epsilon) \tag{13}$$

where $X\beta = \beta x_i + \psi z_i$. x_i is an indicator of whether subject i receives the intervention and z_i is an indicator of whether subject i receives any spillover effect.

3.2 Simulation Procedure

The algorithm implemented to perform the simulations for this study is illustrated in the following section. The simulation is performed over all potential cluster treatment combinations (intervention/control) for each of the 2x4, 3x3, and 3x4 cluster spatial grid geometries. The simulation procedure is as follows:

1. Define the spatial grid geometry & set of parameters for simulation scenario (α, β, ψ, ρ)
2. List out all cluster treatment combinations for specified spatial grid (70 combinations for 2x4, 126 combinations for 3x3, 924 combinations for 3x4
3. Construct the spatial grid object for the corresponding number of clusters (8, 9, 12) and define the rook neighbors for each cluster in the spatial region.
4. Identify cluster treatment combinations for the specified spatial grid object that satisfy Block Stratified Sampling (BSS) conditions (2x4: 2, 3x3: 6, 3x4: 2).
5. For each combination in the set of all cluster treatment combinations for a given spatial grid:
   (a) Sample 20 subjects per cluster over the entire space of the grid (2x4: 160 subjects, 3x3: 180 subjects, 3x4: 240 subjects).

(b) Assign each subject to its corresponding cluster and to the intervention or control treatment based on the treatment assignment of the specific cluster.

(c) Identify the rook neighbors of each subject and whether each subject is exposed to a binary spillover effect based on the treatment assignment of neighboring clusters and selected spillover application type as described in Section 2.3.3.

(d) Generate the response value for each subject with consideration for spatial dependence via Spatial Autoregressive (SAR) modeling.

(e) Estimate and retain parameter values for $\alpha, \beta, \psi,$ and ρ with the Spatial simultaneous autoregressive lag model ('lagsarlm()') via maximum likelihood estimation.

(f) Iterate parts (a)-(e) of Step 5 N=50 times for each cluster treatment combination for the 2x4 and 3x3 spatial grids and N=10 times for each cluster treatment combination for the 3x4 spatial grid.

6. Aggregate parameter estimates for each iteration of each cluster treatment combination for a specified spatial grid

7. For each parameter estimate (α, β, ψ, ρ), compute the bias of the mean of the estimates relative to the true value, variance of the estimates, and Mean Squared Error (MSE) of the estimates such that these metrics are computed for each cluster treatment combination.

8. Repeat Steps 1-5 of the simulation procedure for each spatial grid setup (2x4, 3x3, 3x4).

Figure 4 presents an example of the 8 cluster 2x4 spatial grid object. 160 subjects (20 subjects for each cluster) with uniformly distributed coordinates are sampled over the area of the spatial object and colored according to their cluster membership. Subject counts are not necessarily identical between any pair of clusters as the sample of subjects is considered over the whole space. Points in any cluster on the grid are subject to the selected treatment (intervention/control) and equally experience any potential spillover along with all other points that are members of that same cluster. Similar illustrative figures of the sampled observations over the spatial grid could be constructed and displayed for the other spatial geometries under consideration in the simulation study.

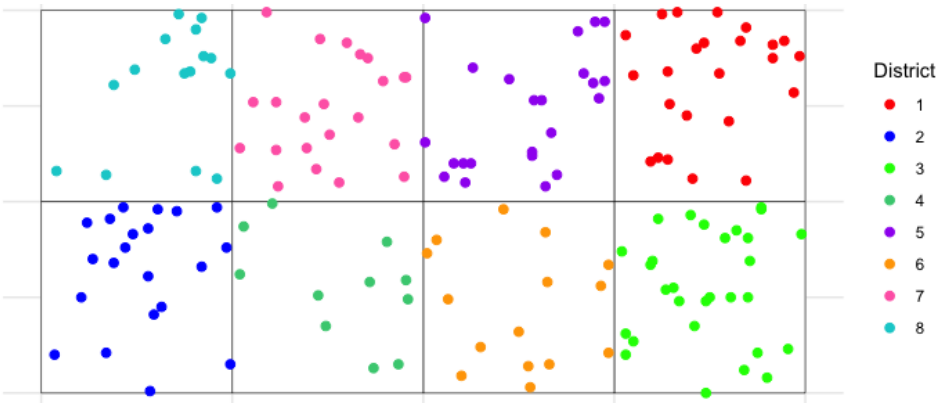

Figure 4: Example of simulated observation coordinates (n=160) on a 2x4 grid of clusters

3.3 Simulation Evaluation Metrics

Estimates for the baseline effect (α), intervention effect (β), spillover effect (ψ), and spatial correlation parameter (ρ) were collected for each iteration of the simulation procedure for every treatment assignment combination. Considering the initial fixed simulation parameters, the following metrics are computed for estimation of the intervention effect for the evaluation of each set of simulation conditions.

Bias: a measure of how different the estimated parameter value is from the true parameter value. Small bias means that the parameter estimate is quite close to the true value.

Variance: a measure of the spread of a set of estimates from multiple simulation iterations. Smaller variance indicates that the estimates are tightly grouped together and large variance means that the estimates are spread out.

Mean Squared Error (MSE): measures the average squared difference between each estimate and the true value of a parameter. Mean Squared Error combines variance to understand the consistency of a set of estimates as well as bias to understand any trend in deviation from the true value into a measure of estimation accuracy.

Mean Squared Error is the primary metric of consideration to understand the efficient estimation of the intervention effect under all simulation settings. The following subsections detail the findings from the simulation study, focusing on the Mean Squared Error (MSE), standard deviation (SD), and bias

of the intervention effect (β) estimates under various conditions (simulation scenarios). The analysis compares Simple Random Sampling (SRS) and Block Stratified Sampling (BSS) across different spatial grid configurations, spillover effect magnitudes, spatial dependence levels, and spillover application types.

3.4 Intervention Effect Estimates (2x4 grid)

For the 2×4 spatial grid, the simulation results for the intervention effect estimates are presented in Table 1 and visually represented in Figure 5.

When spillover was applied to Control Clusters Only (TrtNoSpill), SRS generally exhibited higher MSE compared to BSS, particularly as the spillover effect (ψ) increased. For instance, with $\rho = 0$ (no spatial dependence), SRS MSE ranged from 0.173 ($\psi = 0.5$) to 0.441 ($\psi = 0.8$), while BSS maintained very low MSE values (e.g., 0.008 for $\psi = 0.5$). Similarly, with $\rho = 0.01$ (spatial dependence), SRS MSE increased from 0.304 ($\psi = 0.5$) to 0.573 ($\psi = 0.8$), whereas BSS showed MSE values ranging from 0.165 to 0.569. BSS consistently demonstrated smaller standard deviations compared to SRS under this spillover application type, indicating greater precision. SRS estimates generally showed negative bias, which increased in magnitude with higher ψ.

When spillover was applied to Intervention & Control Clusters (TrtSpill), SRS generally showed lower MSE compared to BSS, especially in the absence of spatial dependence ($\rho = 0$). For $\rho = 0$, SRS MSE ranged from 0.008 ($\psi = 0.5$) to 0.019 ($\psi = 0.8$), while BSS MSE was consistently higher (e.g., 0.262 for $\psi = 0.5$). With spatial dependence ($\rho = 0.01$), SRS still maintained lower MSE (0.095 to 0.162) compared to BSS (0.165 to 0.569). SRS also exhibited smaller standard deviations and bias magnitudes under this condition, suggesting better overall accuracy and precision.

Across both spillover application types for the 2×4 grid, the presence of spatial dependence ($\rho = 0.01$) consistently led to higher MSE values for both sampling methods compared to scenarios without spatial dependence ($\rho = 0$).

3.5 Intervention Effect Estimates (3x3 grid)

The results for the 3×3 spatial grid are summarized in Table 2 and Figure 6.

For the Control Clusters Only (TrtNoSpill) spillover application, BSS consistently yielded significantly lower MSE, SD, and bias compared to SRS. For instance, with $\rho = 0$, BSS MSE was minimal (e.g., 0.001

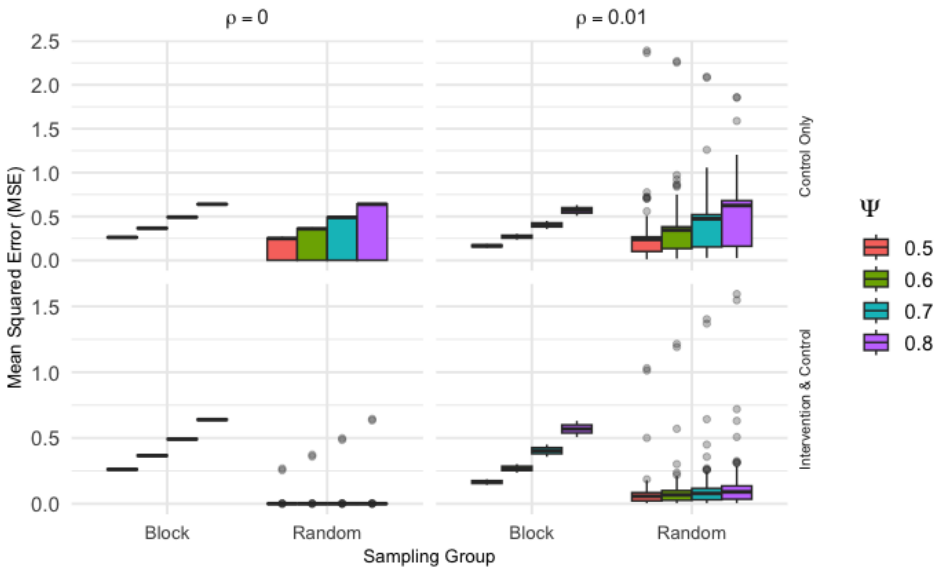

Figure 5: Intervention effect MSE for 2x4 cluster grid grouped by spillover effect magnitude, spatial dependence, sampling design, and spillover application type

for $\psi = 0.5$), while SRS MSE ranged from 0.092 to 0.235. This pattern held true with spatial dependence ($\rho = 0.01$), where BSS MSE remained very low (e.g., 0.004 for $\psi = 0.5$), contrasting with SRS MSE values ranging from 0.259 to 0.372. SRS estimates consistently showed negative bias, increasing with ψ, while BSS bias was negligible.

In the Intervention & Control Clusters (TrtSpill) spillover scenario, BSS continued to demonstrate superior performance with substantially lower MSE, SD, and bias compared to SRS. For $\rho = 0$, BSS MSE was consistently around 0.0004, whereas SRS MSE ranged from 0.002 to 0.005. With $\rho = 0.01$, BSS MSE was also consistently very low (e.g., 0.039 for $\psi = 0.5$), while SRS MSE ranged from 0.103 to 0.168.

Similar to the 2×4 grid, the presence of spatial dependence ($\rho = 0.01$) generally increased MSE for both sampling methods across both spillover application types for the 3×3 grid.

3.6 Intervention Effect Estimates (3x4 grid)

Table 3 and Figure 7 present the intervention effect estimates for the 3×4 spatial grid.

Under the Control Clusters Only (TrtNoSpill) spillover condition, BSS consistently outperformed SRS in terms of MSE, SD, and bias. With $\rho = 0$, BSS MSE was extremely low (e.g., 0.0004 for $\psi = 0.5$), while SRS MSE ranged from 0.034 to 0.219. This trend persisted with $\rho = 0.01$, where BSS MSE remained low

Spillover Application	**Spatial Dependence (ρ)**	**Spillover Effect (ψ)**	**Simple Random Sampling**			**Block Stratified Sampling**		
			MSE	SD ($\times$10)	Bias	MSE	SD ($\times$10)	Bias
Control Only Spillover	Yes (0.01)	0.5	0.304	4.053	-0.270	0.165	0.294	-0.349
		0.6	0.382	4.062	-0.342	0.269	0.462	-0.471
		0.7	0.471	4.143	-0.414	0.404	0.653	-0.594
		0.8	0.573	4.390	-0.486	0.569	0.867	-0.716
	No (0)	0.5	0.173	1.170	-0.343	0.262	0.095	-0.496
		0.6	0.248	1.683	-0.411	0.366	0.100	-0.597
		0.7	0.338	2.291	-0.480	0.492	0.095	-0.697
		0.8	0.441	2.994	-0.548	0.639	0.081	-0.798
Intervention & Control Spillover	Yes (0.01)	0.5	0.095	1.762	0.031	0.165	0.294	-0.349
		0.6	0.115	2.082	0.030	0.269	0.462	-0.471
		0.7	0.138	2.423	0.029	0.404	0.653	-0.594
		0.8	0.162	2.787	0.027	0.569	0.867	-0.716
	No (0)	0.5	0.008	0.438	-0.014	0.262	0.095	-0.496
		0.6	0.011	0.614	-0.017	0.366	0.100	-0.597
		0.7	0.015	0.825	-0.020	0.492	0.095	-0.697
		0.8	0.019	1.072	-0.023	0.639	0.081	-0.798

Table 1: SRS vs. BSS: Intervention Effect Error Estimates - 2x4 Spatial Grid Setting

(e.g., 0.035 for $\psi = 0.5$), compared to SRS MSE ranging from 0.120 to 0.204. SRS estimates generally showed a small positive or negative bias, while BSS bias was consistently negligible.

For the Intervention & Control Clusters (TrtSpill) spillover application, BSS again showed superior performance with lower MSE, SD, and bias. With $\rho = 0$, BSS MSE was consistently around 0.0003, while SRS MSE ranged from 0.0007 to 0.029. With $\rho = 0.01$, BSS MSE was also low (e.g., 0.223 for

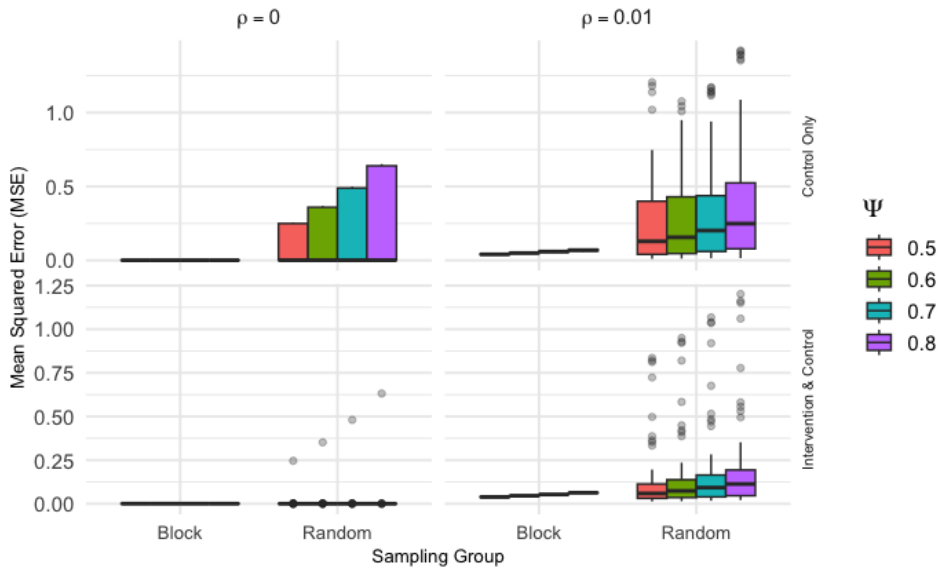

Figure 6: Intervention effect MSE for 3x3 cluster grid grouped by spillover effect magnitude, spatial dependence, sampling design, and spillover application type

$\psi = 0.5$), while SRS MSE ranged from 0.112 to 0.190.

Across all grid sizes and conditions, increasing the spillover effect magnitude (ψ) generally led to an increase in MSE for both sampling methods. The presence of spatial dependence ($\rho = 0.01$) consistently resulted in higher MSE for both SRS and BSS compared to scenarios without spatial dependence ($\rho = 0$). While SRS sometimes showed lower MSE when spillover was applied to both intervention and control clusters, BSS consistently demonstrated lower variance and often lower bias, particularly when spillover was restricted to control clusters only. This highlights that BSS provides more consistent estimates, even if SRS might occasionally yield a lower average error depending on the specific spillover dynamics.

4 Application

The findings from our simulation study provide relevant insights for investigators designing CRTs in real-world applications where spatial dependence and spillover effects are present and must be accounted for. In this case, a pair of studies in development that we consider involve the North Carolina Department of Adult Correction. The structure of correctional systems, particularly community supervision, often involves geographically clustered units where staff from different clusters may interact and regional similarities can influence outcomes. The investigators leading these studies desire recommendations for

Spillover Application	Spatial Dependence (ρ)	Spillover Effect (ψ)	Simple Random Sampling			Block Stratified Sampling		
			MSE	SD (×10)	Bias	MSE	SD (×10)	Bias
Control Only Spillover	Yes (0.01)	0.5	0.234	2.593	-0.054	0.041	0.044	-0.012
		0.6	0.276	2.860	-0.092	0.050	0.080	-0.018
		0.7	0.322	3.284	-0.131	0.059	0.114	-0.022
		0.8	0.372	3.857	-0.170	0.069	0.145	-0.025
	No (0)	0.5	0.092	1.207	-0.182	0.001	0.002	0.003
		0.6	0.132	1.740	-0.219	0.001	0.002	0.003
		0.7	0.180	2.370	-0.255	0.001	0.002	0.003
		0.8	0.235	3.099	-0.292	0.001	0.002	0.003
Intervention & Control Spillover	Yes (0.01)	0.5	0.103	1.480	0.071	0.039	0.022	-0.019
		0.6	0.122	1.688	0.082	0.046	0.027	-0.020
		0.7	0.143	1.905	0.092	0.054	0.032	-0.021
		0.8	0.168	2.149	0.101	0.063	0.007	-0.019
	No (0)	0.5	0.002	0.219	-0.004	0.001	0.004	0.004
		0.6	0.003	0.313	-0.004	0.001	0.004	0.004
		0.7	0.004	0.428	-0.005	0.001	0.004	0.004
		0.8	0.005	0.562	-0.006	0.001	0.004	0.004

Table 2: SRS vs. BSS: Intervention Effect Error Estimates - 3x3 Spatial Grid Setting

treatment assignment methodology regarding the spatially clustered nature of their design.

4.1 Background of Probation in North Carolina

The North Carolina Department of Adult Correction Division of Community Supervision (NC DCS) provides a clear example of a spatially structured organization, making it an ideal context for applying

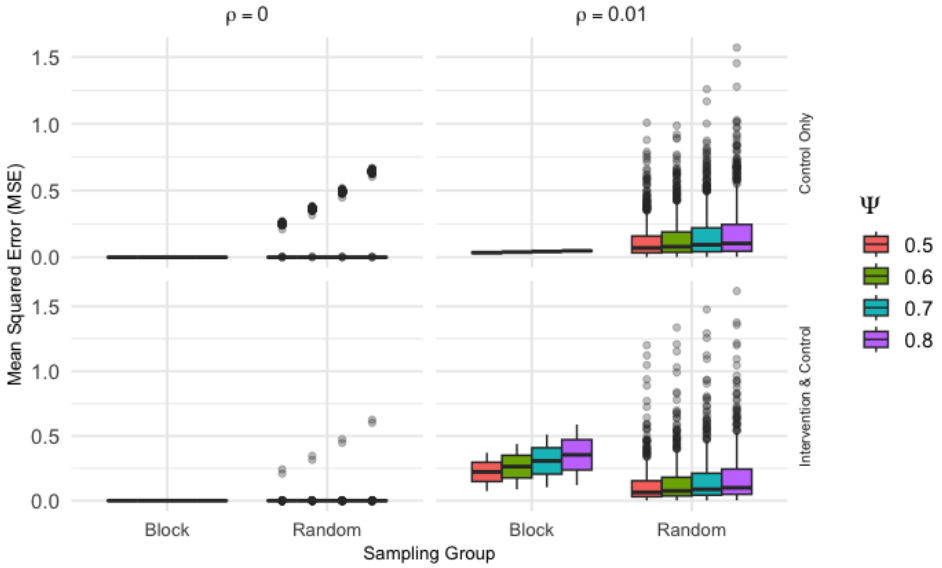

Figure 7: Intervention effect MSE for 3x4 cluster grid grouped by spillover effect magnitude, spatial dependence, sampling design, and spillover application type

our research. The NC DCS is organized into four geographic Divisions numbered 1 through 4 from East to West, which are subdivided into 30 Districts that make up the entirety of the state. Each district is composed of one or more counties. If there is more than one county in the district, the counties are adjacent, but some districts are composed of only a singular county. In these districts, Probation and Parole Officers (PPOs) supervise individuals at the county level and are themselves managed by Chief Probation/Parole Officers (CPPOs) and Judicial District Managers (JDMs) who oversee districts. Each division (composed of multiple districts) is managed by a Judicial Division Director (JD Dir) as well. A map of the segmentation of each division and district is presented in Figure 8.

This geographic and administrative structure means that Chiefs and Judicial District Managers (JDMs) often span multiple units and counties, creating a direct channel for spillover effects among regions under their supervision. For instance, if leadership views an intervention in one county favorably, they may implement its components in other counties under their jurisdiction, even if those areas were not assigned the intervention. Furthermore, adjacent counties within the same region often share demographic characteristics, labor market conditions, and similar judicial cultures, as well as the same managed care organizations and treatment providers, leading to spatial dependence in probation outcomes. The choice of treatment assignment sampling strategy in this context is non-trivial when considering potential options of either simple random sampling (SRS) or block-stratified sampling (BSS).

Spillover Application	**Spatial Dependence (ρ)**	**Spillover Effect (ψ)**	**Simple Random Sampling**			**Block Stratified Sampling**		
			MSE	SD (×100)	Bias	MSE (×10)	SD (×100)	Bias
Control Only Spillover	Yes (0.01)	0.5	0.120	13.244	0.033	0.346	0.963	-0.152
		0.6	0.138	14.807	0.021	0.388	0.505	-0.167
		0.7	0.159	17.166	0.009	0.435	0.011	-0.180
		0.8	0.183	20.387	-0.001	0.484	0.561	-0.193
	No (0)	0.5	0.034	8.554	-0.068	0.004	0.022	-0.006
		0.6	0.049	12.318	-0.081	0.004	0.022	-0.006
		0.7	0.067	16.769	-0.094	0.004	0.020	-0.006
		0.8	0.087	21.912	-0.108	0.003	0.016	-0.006
Intervention & Control Spillover	Yes (0.01)	0.5	0.112	12.873	0.022	2.229	20.995	-0.415
		0.6	0.131	14.762	0.021	2.634	24.750	-0.452
		0.7	0.151	16.817	0.020	3.074	28.774	-0.488
		0.8	0.173	19.046	0.018	3.548	33.060	-0.525
	No (0)	0.5	0.001	1.051	-0.001	0.003	0.004	-0.003
		0.6	0.001	1.542	-0.001	0.003	0.004	-0.002
		0.7	0.001	2.143	-0.002	0.003	0.005	-0.002
		0.8	0.002	2.853	-0.002	0.003	0.005	-0.002

Table 3: SRS vs. BSS: Intervention Effect Error Estimates - 3x4 Spatial Grid Setting

Selection of the appropriate treatment sampling strategy may yield beneficial downstream outcomes when considering the error of intervention & spillover estimates.

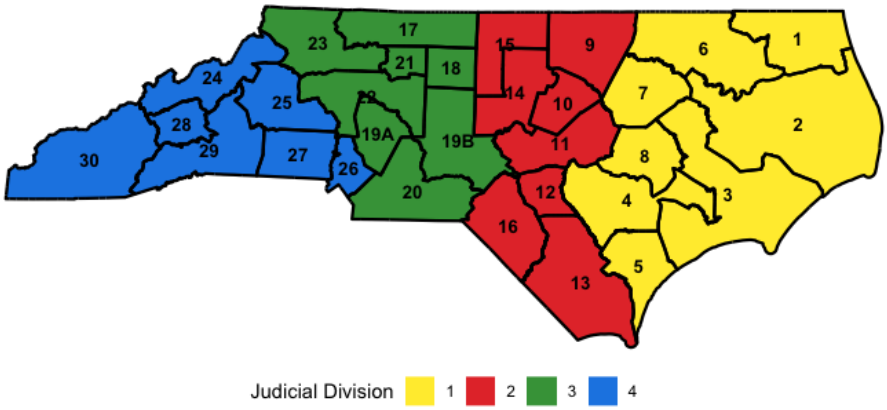

Figure 8: The North Carolina Department of Adult Correction Division of Community Supervision map is segmented into 30 judicial districts that are grouped into 4 judicial divisions.

4.2 Case 1: Trauma-Informed and Engaged Supervision

One proposed study involves implementing a trauma-informed supervision model for PPOs managing specialty caseloads like domestic violence, sex offender, or gang affiliated cases. The intervention consists of intensive training and follow-up sessions delivered to cohorts of PPOs and CPPOs. The primary outcomes would be changes in officer knowledge and attitudes, as well as impacts on individuals under their supervision. The key design decision for this study is the unit of randomization. Randomizing at the PPO level is not feasible due to the high likelihood of spillover within a county office. Therefore, a cluster randomized design at the county or district level is necessary. This cluster randomization decision is also supported by agency-related logistics in addition to considerations for spillover. Original plans for a similar study, acknowledging the risk of contamination between adjacent units, considered randomization at the district level specifically to minimize spillover. This scenario directly reflects the conditions modeled in our simulation, where the effectiveness of an intervention in one cluster may influence outcomes in neighboring control clusters where an intervention is not applied.

4.3 Case 2: Testing an implementation strategy for specialty mental health probation

Another study proposal aims to test a strategy for improving the implementation of Specialty Mental Health Probation (SMHP). The intervention involves enhancing collaboration between PPOs and community-based behavioral health service providers through cross-training sessions and joint treatment team meetings. The primary outcome is improved collaboration, with secondary outcomes related to treatment engagement for individuals on probation. This case presents a more complex spatial challenge. In addition to the geographic proximity of probation districts, the behavioral health system in North Carolina is managed by four large Managed Care Organizations (MCOs), each responsible for a different coverage area. Some service provider agencies are large, operating in multiple counties and potentially across different MCO regions. This creates multiple, overlapping layers of potential spillover. Figure 9 presents a map of which counties in the state that each MCO operates. An intervention implemented in one county could be adopted by a provider and informally carried into a neighboring control county, or an MCO could promote a similar initiative, confounding the study's results. This complex spatial linkage underscores the importance of a robust treatment assignment strategy in order to efficiently estimate the impact of the intervention.

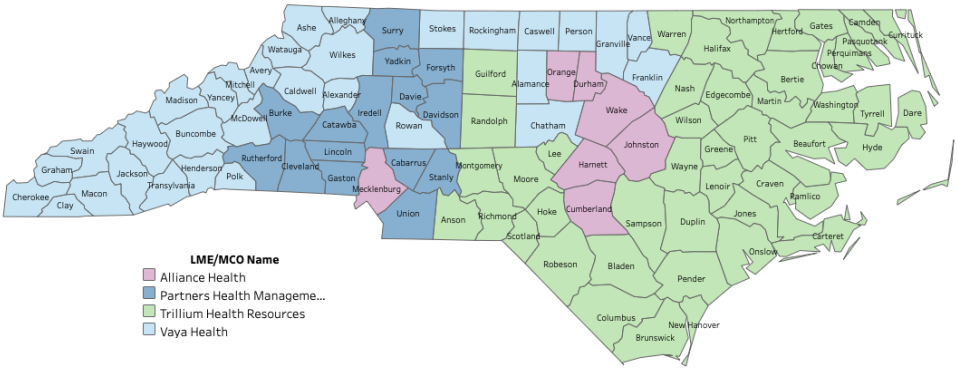

Figure 9: Map of Local Management Entity/Managed Care Organizations (LME/MCOs) for each county in the state of North Carolina. There are 4 LME/MCOs of which one operates in each of the 100 counties of the state.

4.4 Recommendations for treatment assignment

The results of our simulation study indicate a clear trade-off: we previously concluded that Simple Random Sampling (SRS) reduces the average mean squared error (MSE) across many randomizations (intervention/control combinations), while Block Stratified Sampling (BSS) reduces the variance of the error although it is not minimized. Put simply, BSS is less likely to produce a "perfect" assignment by chance but is also robust against a "worst-case" assignment that could render the study results invalid. For both case studies presented, Block Stratified Sampling (BSS) is the recommended approach.

In Case 1 (Trauma-Informed Supervision), a simple random assignment of districts could, by chance, result in a geographic imbalance, concentrating treated districts in one part of the state. This would confound the intervention effect with regional characteristics. A more robust design would be to stratify by the four state divisions and randomize districts to intervention and control conditions within each division with a block stratification strategy. This ensures geographic balance and leverages the finding that BSS minimizes estimation variance, providing a more stable and defensible estimate of the intervention effect.

In Case 2 (SMHP Implementation), the rationale for BSS is even stronger due to the complex potential for spillover. A simple random assignment of counties would ignore the significant structural influence of the MCOs. The most effective design would be to stratify by MCO coverage area, randomizing counties to the implementation strategy (intervention) or control condition within each MCO. This approach directly controls for the significant confounding variable of MCO-level policies and provider networks. However, some challenges with this MCO stratification do persist. For example, it's possible that a county is selected for the intervention and is supervised by a CPPO or even a judicial district manager that oversees a control group county. Furthermore, multiple service providers may serve a given county in the coverage area of an MCO and it is possible that multiple service providers begin team meetings. Another consideration is that some service providers are very large and have offices around the state and contract with more than one MCO. Based on this, we must consider potential spillover if after participating in a treatment team meeting, a service provider intends initiate those practices in another part of the state. Given the high risk of contamination, the robustness of BSS in minimizing error variance makes it the most prudent choice for achieving a reliable effect estimate. While SRS might achieve a lower MSE (intervention error estimate) in theory, the risk of a single, poorly distributed

randomization in a complex, real-world setting outweighs the potential benefit. Therefore, it is wise to proceed with a sampling strategy that aims to manage estimation error with minimal variation.

A representation of this Block Stratification sampling strategy where we "block" treatment and control districts together with an effort to isolate districts of a particular intervention status from each other is shown in Figure 10. Note that across the entire map of the state and specifically in Judicial Division 1 (easternmost region) that when possible districts of the same intervention status are isolated from each other. When districts share a border with multiple other districts, this sampling method ensures that treatment districts are isolated from other treatment districts as best as possible.

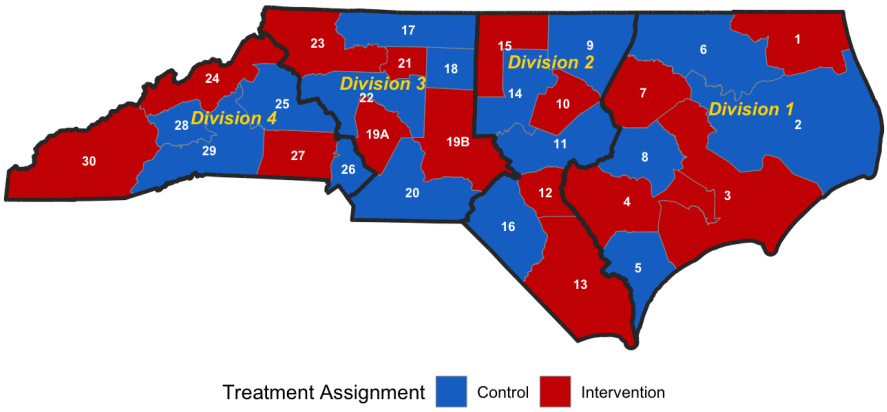

Figure 10: Map of North Carolina with example block stratified sampling intervention assignments for each judicial district under consideration. Districts are assigned to either treatment (red) or control (blue).

5 Discussion

5.1 Relevance of results

This simulation study provides insights into the efficiency of treatment assignment methodologies in spatial cluster randomized trials (CRTs) when considering spillover effects and spatial dependence. A primary contribution of this work is the direct comparison between Simple Random Sampling (SRS) and Block Stratified Sampling (BSS) across various scenarios, offering practical guidance for investigators

designing such trials.

Our findings show a clear relationship between the magnitude of spillover effects and the estimation error of the intervention effect. As the spillover effect (ψ) increases, the Mean Squared Error (MSE) for the intervention effect (β) generally rises, regardless of the sampling method or the presence of spatial dependence. This highlights the importance of accounting for spillover effects, as their increasing influence can diminish the accuracy of intervention effect estimates. The results indicate that randomly selected treatment assignments lead to reduced average error (lower bias) when estimating intervention effects, while block stratified treatment assignments consistently result in reduced variation (lower variance) among a series of estimates. This suggests a trade-off: SRS provides a more accurate central estimate on average, whereas BSS offers greater precision and consistency across repeated trials, potentially with higher bias.

The study also clarifies the impact of spatial dependence (ρ) on intervention effect estimation error. The presence of even a small degree of spatial dependence (e.g., $\rho = 0.01$) generally increased MSE for both sampling methods. This indicates that spatial correlation among observations can impede the accurate estimation of intervention effects if not managed during the design phase. The observed patterns across different spatial grid dimensions (2×4, 3×3, and 3×4) support these conclusions, demonstrating the consistency of the findings across varying spatial configurations.

The investigation into cluster spillover restrictions (spillover to control clusters only vs. spillover to both control and intervention clusters) also identified important differences. When spillover was restricted to control clusters only, the MSE for the intervention effect was higher for SRS compared to BSS, particularly with increasing spillover magnitude. Conversely, when spillover was permitted for both control and intervention clusters, SRS often exhibited lower MSE, especially when spatial dependence was absent. This understanding of spillover application types provides investigators with a more precise basis for selecting a sampling strategy that aligns with their specific hypotheses regarding how an intervention's effects might propagate.

It is also insightful to note that for each simulation grid setup (2×4, 3×3, and 3×4), there are 7 different treatment assignment combinations that are the overall "optimal" combination for at least one of the simulation scenarios in terms of MSE of the intervention effect estimate. It appears that the optimal treatment combination shifts as the components of each simulation scenario are adjusted and

there is no consistent treatment combination from the Simple Random Sampling strategy that is optimal. This serves as further evidence to suggest that it is prudent to recommend the Block Stratified Sampling strategy because it provides results that consistently exhibit minimal MSE of the intervention effect estimate although the BSS treatment combinations are never the absolute best case in terms of MSE for any given simulation scenario. By implementing a BSS strategy, investigators will consistently obtain strong estimates of the intervention effect (although not the best for any specific single case) under the conditions of the spatial autoregressive (SAR) model.

In summary, this study offers a comparison of two opposing sampling methodologies in spatial CRTs. The patterns observed regarding spillover magnitude, spatial dependence, and spillover application types provide practical information for investigators. They must consider the goal of an unbiased average estimate (favored by SRS) versus the need for consistent, low-variance estimates (achieved with BSS), especially when designing studies where spillover effects are expected.

5.2 Limitations

While this simulation study provides valuable insights, it is important to acknowledge a few key limitations that may affect the generalizability and scope of its findings. The study encountered issues related to multicollinearity between the direct (intervention) and indirect (spillover) effects. In some spatial configurations and with specific spillover magnitudes, high correlation between the x_i (intervention indicator) and z_i (spillover indicator) variables in the spatial autoregressive model (Equation 1) could lead to instability in parameter estimates. While the SAR model addresses spatial relationships, extreme collinearity can still impact the precision of individual coefficient estimates, potentially affecting MSE calculations. This issue in separating highly correlated effects is known in spatial causal inference and requires careful consideration in applied settings. It is also necessary to note that the expanded simulation space quickly becomes computationally infeasible. The current study explored a specific range of spillover effect sizes, spatial dependence parameters, and grid dimensions. However, a more extensive exploration of additional parameters (e.g., different baseline effects, varying intervention effect sizes, or a wider range of spatial dependence values) or more complex spatial structures would require significantly greater computational resources. This limitation restricted the number of scenarios that could be thoroughly investigated within the current framework.

An additional notable limitation is the rudimentary spillover mechanism (binary effect) applied. This simulation study modeled spillover as a fixed, binary event, meaning a unit either experienced the full potential spillover effect or none, based solely on contiguity (Rook neighbors). This approach does not account for more realistic scenarios where spillover effects may decay with distance, or where cumulative exposure from multiple treated neighbors could lead to additive or non-linear effects. The absence of more sophisticated distance-based decay functions (e.g., linear, categorical, exponential, or Gaussian decay) or mechanisms for additive spillover limits the direct applicability of these specific quantitative results to real-world situations where such nuanced spillover dynamics are likely. The current model also did not consider higher-order neighbors beyond immediate (Rook) contiguity. The study also utilized homogeneous cluster sizes and regular grid shapes. Real-world clusters often exhibit irregular shapes and varying sizes, which could influence spillover dynamics and the efficiency of different sampling strategies. The current findings may not fully extend to such heterogeneous spatial arrangements. Furthermore, the study assumes a fixed treatment allocation ratio (primarily 1:1 or near 1:1). Varying the proportion of intervention to control clusters could impact the statistical power and bias of intervention effect estimates, especially in the presence of spillovers. Finally, it is crucial to note that the results obtained and recommendation for a Block Stratified Sampling (BSS) strategy hold under the conditions of the spatial autoregressive (SAR) model and it cannot be guaranteed that the apparent benefits of BSS persist under the conditions of an alternative spatial model.

5.3 Future Considerations

Building on this study's findings, several areas for future research can further improve our understanding of sampling design in spatial cluster randomized trials. A key next step involves incorporating additional spillover structures beyond the binary mechanism. Future simulations should explore various distance-based decay functions, such as linear, categorical/step-wise, exponential, or Gaussian decay. This would provide a more realistic representation of how intervention effects attenuate over space and allow for the identification of optimal sampling strategies under different decay profiles. For example, an intervention with rapid exponential decay might benefit from a different sampling approach than one with a more gradual linear decay.

Furthermore, the current model's assumption of non-additive spillover effects could be relaxed. Fu-

ture research should consider the implications of "additive" spillover, where being a neighbor to multiple intervention cells allows for a greater cumulative spillover effect. This would reflect scenarios where the intensity of exposure to the intervention's indirect effects scales with the number or proximity of treated neighbors. Understanding how sampling methods perform under such additive spillover mechanisms is important for interventions where collective exposure drives indirect effects. Exploring other neighbor structures beyond Rook contiguity is also warranted. Investigating Queen neighbors (sharing an edge or vertex) or higher-order neighbors (neighbors of neighbors) would provide a more complete understanding of how different definitions of "proximity" influence spillover propagation and, consequently, the performance of various sampling designs.

Additionally, future work could investigate the impact of "prior" weighting of intervention based on cluster subject density. In real-world settings, clusters rarely have uniform population density. Incorporating variable subject densities within clusters and exploring how this influences spillover effects and optimal sampling strategies would add realism to the simulations. An extension of the simulation parameters to cover a broader range of values for spillover effects, spatial dependence, and grid dimensions would allow for a more comprehensive mapping of the parameter space and a more robust set of recommendations for practitioners. This would enable a deeper understanding of the conditions under which each sampling method (SRS vs. BSS) is most advantageous, leading to more efficient and reliable designs for spatial cluster randomized trials.

Future research could also investigate the robustness of findings to model misspecification. This includes evaluating how different sampling strategies perform when alternative spatial models (e.g., Spatial Error Model, Spatial Durbin Model) are used for analysis, or when the true underlying spatial process deviates from the assumed SAR model. Finally, exploring optimal allocation ratios of intervention to control clusters in the presence of spillover effects and spatial dependence could provide valuable insights for maximizing statistical power or minimizing estimation error. Incorporating cost-effectiveness analysis into the design framework would also be beneficial, as different sampling strategies may have varying logistical and financial implications in real-world implementation.

Authors' Contributions

A.W. wrote the manuscript text and prepared figures 1-7 & 10. T.V. provided figures 8 & 9 and provided advisement for the text of section 4 of the manuscript. All authors reviewed the manuscript.

Funding

Dr. Van Deinse is supported by the National Institutes of Mental Health (K01MH129619). The trauma-informed supervision approached referenced in this manuscript was funded by the Bureau of Justice Assistance 15PBJA-24-GG-01896-PRJH and awarded to the North Carolina Department of Adult Correction.

Data Availability

All data supporting the findings of this study are generated via simulation for the included simulation study. Summarized findings are included in tables 1-3. The simulation data are available upon request and with permission of the authors.

Declarations

Ethics, Consent to Participate, and Consent to Publish

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

[1] Stuart J. Pocock. *Clinical trials: a practical approach.* A Wiley medical publication. Wiley, Chichester [West Sussex] ; New York, 1983.

[2] Peter G. Smith, Richard H. Morrow, and David A. Ross, editors. *Field Trials of Health Interventions: A Toolbox.* Wellcome Trust–Funded Monographs and Book Chapters. OUP Oxford, Oxford (UK), 3rd edition, 2015.

[3] Richard J. Hayes and Lawrence H. Moulton. *Cluster Randomised Trials, Second Edition*. Chapman and Hall/CRC, July 2017.

[4] Allan Donner and Neil Klar. *Design and analysis of cluster randomization trials in health research*. Arnold, London, 1. publ edition, 2000.

[5] Michael G Hudgens and M. Elizabeth Halloran. Toward Causal Inference With Interference. *Journal of the American Statistical Association*, 103(482):832–842, June 2008.

[6] R J Hayes, N De Alexander, S Bennett, and S N Cousens. Design and analysis issues in cluster-randomized trials of interventions against infectious diseases. *Statistical Methods in Medical Research*, 9(2):95–116, April 2000.

[7] Karim Anaya-Izquierdo and Neal Alexander. Spatial regression and spillover effects in cluster randomized trials with count outcomes. *Biometrics*, 77(2):490–505, June 2021.

[8] Christopher Jarvis, Gian Luca Di Tanna, Daniel Lewis, Neal Alexander, and W. John Edmunds. Spatial analysis of cluster randomised trials: a systematic review of analysis methods. *Emerging Themes in Epidemiology*, 14(1):12, September 2017.

[9] F. N. Binka, F. Indome, and T. Smith. Impact of spatial distribution of permethrin-impregnated bed nets on child mortality in rural northern Ghana. *The American Journal of Tropical Medicine and Hygiene*, 59(1):80–85, July 1998.

[10] John E. Gimnig, Margarette S. Kolczak, Allen W. Hightower, John M. Vulule, Erik Schoute, Luna Kamau, Penelope A. Phillips-Howard, Feiko O. ter Kuile, Bernard L. Nahlen, and William A. Hawley. Effect of permethrin-treated bed nets on the spatial distribution of malaria vectors in western Kenya. *The American Journal of Tropical Medicine and Hygiene*, 68(4 Suppl):115–120, April 2003.

[11] William A. Hawley, Penelope A. Phillips-Howard, Feiko O. ter Kuile, Dianne J. Terlouw, John M. Vulule, Maurice Ombok, Bernard L. Nahlen, John E. Gimnig, Simon K. Kariuki, Margarette S. Kolczak, and Allen W. Hightower. Community-wide effects of permethrin-treated bed nets on child mortality and malaria morbidity in western Kenya. *The American Journal of Tropical Medicine and Hygiene*, 68(4 Suppl):121–127, April 2003.

[12] Edward Miguel and Michael Kremer. Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities. *Econometrica*, 72(1):159–217, January 2004.

[13] W. R. Tobler. A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, 46:234, June 1970.

[14] Arthur Getis. A History of the Concept of Spatial Autocorrelation: A Geographer's Perspective. *Geographical Analysis*, 40(3):297–309, July 2008.

[15] Noel A. C. Cressie. *Statistics for Spatial Data*. Wiley Series in Probability and Statistics. Wiley, 1 edition, September 1993.

[16] Audrey Lenhart, Nicolas Orelus, Rachael Maskill, Neal Alexander, Thomas Streit, and P. J. McCall. Insecticide-treated bednets to control dengue vectors: preliminary evidence from a controlled trial in Haiti. *Tropical Medicine & International Health*, 13(1):56–67, January 2008.

[17] D. L. Chao, J. K. Park, F. Marks, R. L. Ochiai, I. M. Longini, and M. E. Halloran. The contribution of neighbours to an individual's risk of typhoid outcome. *Epidemiology and Infection*, 143(16):3520–3527, December 2015.

[18] Neal D Alexander, Rana A Moyeed, Phil J Hyun, Zachary B Dimber, Moses J Bockarie, Julian Stander, Bryan T Grenfell, James W Kazura, and Michael P Alpers. Spatial variation of Anopheles-transmitted Wuchereria bancrofti and Plasmodium falciparuminfection densities in Papua New Guinea. *Filaria Journal*, 2(1):14, September 2003.

[19] Paul Silcocks and Denise Kendrick. Spatial effects should be allowed for in primary care and other community-based cluster RCTS. *Trials*, 11(1):55, December 2010.

[20] Samuel I. Watson and Thomas A. Smith. Design and Analysis of Randomized Trials to Estimate Spatio-Temporally Heterogeneous Treatment Effects. *Journal of the American Statistical Association*, pages 1–11, September 2025.

[21] KyungMann Kim, Frank Bretz, Ying Kuen K. Cheung, and Lisa V. Hampson. *Handbook of Statistical Methods for Randomized Controlled Trials*. Chapman and Hall/CRC, Boca Raton, 1 edition, July 2021.

[22] John N.S. Matthews. *Introduction to Randomized Controlled Clinical Trials*. Chapman and Hall/CRC, 0 edition, June 2006.

[23] David M. Murray. *Design and analysis of group-randomized trials*. Number v. 27 [i.e. 29] in Monographs in epidemiology and biostatistics. Oxford University Press, New York, 1998.

[24] K Hemming, S Eldridge, G Forbes, C Weijer, and M Taljaard. How to design efficient cluster randomised trials. *BMJ*, page j3064, July 2017.

[25] Michael J. Campbell and Stephen John Walters. *How to design, analyse and report cluster randomised trials in medicine and health related research*. John Wiley & Sons, Chichester, West Sussex, 2014.

[26] Georgia Ntani, Hazel Inskip, Clive Osmond, and David Coggon. Consequences of ignoring clustering in linear regression. *BMC Medical Research Methodology*, 21(1):139, December 2021.

[27] A. Stewart Fotheringham and Peter Rogerson. *The Sage handbook of spatial analysis*. The Sage handbook of. Sage, London, 2009.

[28] Roger Bivand. R Packages for Analyzing Spatial Data: A Comparative Case Study with Areal Data. *Geographical Analysis*, 54(3):488–518, July 2022.

[29] Roger S. Bivand and David W. S. Wong. Comparing implementations of global and local indicators of spatial association. *TEST*, 27(3):716–748, September 2018.

[30] Edzer Pebesma and Roger Bivand. *Spatial Data Science: With Applications in R*. Chapman and Hall/CRC, New York, 1 edition, May 2023.

[31] Roger Bivand, Edzer J. Pebesma, and Virgilio Gómez-Rubio. *Applied Spatial Data Analysis with R*. Use R! Springer New York, New York, NY, 2008.

[32] Paula Moraga. *Spatial statistics for data science: theory and practice with R*. Chapman and Hall/CRC Data Science Series. CRC Press, Boca Raton London New York, first edition edition, 2024.

[33] Jade Benjamin-Chung, Jaynal Abedin, David Berger, Ashley Clark, Veronica Jimenez, Eugene Konagaya, Diana Tran, Benjamin F Arnold, Alan E Hubbard, Stephen P Luby, Edward Miguel,

and John M Colford. Spillover effects on health outcomes in low- and middle-income countries: a systematic review. *International Journal of Epidemiology*, 46(4):1251–1276, August 2017.

[34] Jade Benjamin-Chung, Haodong Li, Anna Nguyen, Gabriella Barratt Heitmann, Adam Bennett, Henry Ntuku, Lisa M. Prach, Munyaradzi Tambo, Lindsey Wu, Chris Drakeley, Roly Gosling, Davis Mumbengegwi, Immo Kleinschmidt, Jennifer L. Smith, Alan Hubbard, Mark Van Der Laan, and Michelle S. Hsiang. Extension of efficacy range for targeted malaria-elimination interventions due to spillover effects. *Nature Medicine*, 30(10):2813–2820, October 2024.

[35] Edzer Pebesma. Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, 10(1):439, 2018.

[36] Lance A. Waller. *Applied Spatial Statistics for Public Health Data*. Number v.368 in Wiley Series in Probability and Statistics Ser. John Wiley & Sons, Incorporated, Hoboken, 1st ed edition, 2004.

[37] Aman Ullah and David E. A. Giles. *Handbook of Applied Economic Statistics*. Number v.Vol. 155 in Statistics: a Series of Textbooks and Monographs. Chapman and Hall/CRC, Boca Raton, 1998.

[38] P. A. P. Moran. The Interpretation of Statistical Maps. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 10(2):243–251, July 1948.

[39] P. A. P. Moran. Notes on Continuous Stochastic Phenomena. *Biometrika*, 37(1/2):17, June 1950.

[40] R. C. Geary. The Contiguity Ratio and Statistical Mapping. *The Incorporated Statistician*, 5(3):115, November 1954.

[41] Keith Ord. Estimation Methods for Models of Spatial Interaction. *Journal of the American Statistical Association*, 70(349):120–126, March 1975.

[42] Roger Bivand and Gianfranco Piras. Comparing Implementations of Estimation Methods for Spatial Econometrics. *Journal of Statistical Software*, 63(18), 2015.

[43] Roger Bivand, Jan Hauke, and Tomasz Kossowski. Computing the Jacobian in Gaussian spatial autoregressive models: An illustrated comparison of available methods. *Geographical Analysis*, 45(2):150–179, April 2013.

[44] Luc Anselin. Spatial Econometrics. In Badi H. Baltagi, editor, *A Companion to Theoretical Econometrics*, pages 310–330. Wiley, 1 edition, January 2003.